

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 14:10:20 ; Search time 2707 Seconds
(without alignments)
7994.530 Million cell updates/sec

Title: US-10-027-632-1
Perfect score: 529
Sequence: 1 catgtgtatgtgttagtcgt.....tcttggagagggtcttgag 529

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sta.*
- 12: gb.ey.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sta.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.ey.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	517.6	97.8	117978	9	AC079200	Homo sapi
2	517.6	97.8	167118	9	AC023385	Homo sapi
3	451.6	85.4	63888	2	AC084244	Homo sapi
4	58.4	11.0	341050	3	PFA929357	Plasmid
5	52.4	9.9	251551	3	AB014844	Plasmid
6	50.8	9.6	146570	3	AC117072	Dictyoste
7	50.6	9.6	958	3	DDI1092A	M19468 D.discoideu
8	50.6	9.6	2255	3	DDITUBA	L13999 Dictyosteli
9	50.6	9.6	331039	3	AC116988	Dictyoste
10	50.4	9.5	110000	2	PFMAL8P1_01	Continuation (2 of
11	49.8	9.4	16744	2	EX469924	DX469924 Danio rer
12	49.6	9.4	518	3	AY221171	Dictyoste
13	49.4	9.3	256879	3	AC116982	Dictyoste
14	49.2	9.3	5850	3	DDIDDP2	M55298 Dictyosteli
15	49.2	9.3	5852	3	DDP2PLAS	X51478 Dictyosteli
16	48.8	9.2	302156	3	AZ1622	A21622 Dictyosteli
17	48.8	9.2	302156	3	AC116977	Dictyoste
18	48.6	9.2	611	6	AX187051	AX187051 Sequence
19	48.6	9.2	343050	3	PFA929353	Plasmid
20	48.4	9.1	1052	3	AF020696	AF020696 Trypanoso
21	48.4	9.1	253305	3	PFMAL3P7	AL034559 Plasmid
22	48.2	9.1	39989	2	AC091136	Homo sapi
23	48.2	9.1	146150	2	AC021298	Homo sapi
24	48.2	9.1	155985	9	AC124657	Homo sapi
25	48.2	9.1	181996	2	AC068749	Homo sapi
26	48.2	9.1	256221	2	AL356371	Homo sapi
27	47.8	9.0	36188	3	AC115608	Dictyoste
28	47.8	9.0	54441	3	AC115584	Dictyoste
29	47.8	9.0	152336	2	AC009594	Homo sapi
30	47.8	9.0	162554	9	AC079354	Homo sapi
31	47.8	9.0	182012	9	AL359846	Human DNA
32	47.6	9.0	6457	3	AC115613	Dictyoste
33	47.6	9.0	61052	2	AC123513	Dictyoste
34	47.6	9.0	136240	3	AC117070	Dictyoste
35	47.6	9.0	163443	2	AC006280	Plasmid
36	47.6	9.0	196149	2	AC004709	Plasmid
37	47.6	9.0	250029	3	AE014830	Plasmid
38	47.6	9.0	252650	3	AE014847	Plasmid
39	47.4	9.0	110000	3	PFMAL1P2_1	Continuation (2 of
40	47.4	9.0	113515	6	AX347077	Sequence
41	47.4	9.0	156060	2	AC004153	Plasmid
42	47.4	9.0	176109	9	AC003992	Homo sapi
43	47.4	9.0	176379	9	AC016751	Homo sapi
44	47.4	9.0	250707	3	AE014848	Plasmid
45	47.4	9.0	253305	3	PFMAL3P7	Plasmid
46	47.2	8.9	2992	3	AF012945	Dictyoste
47	47.2	8.9	111882	3	AC115612	Dictyoste
48	47.2	8.9	340552	3	PFA929354	Plasmid
49	47.2	8.9	611	6	AX187051	Sequence
50	47.2	8.9	24479	3	AC114265	Dictyoste
51	47.2	8.9	105682	3	AC116957_3	Continuation (4 of
52	46.8	8.8	2245	3	AE016009_02	Continuation (3 of
53	46.8	8.8	12029	3	AE001427	Continuation (3 of
54	46.8	8.8	110000	2	AC140024	Medicago
55	46.8	8.8	118587	2	AC140024	Medicago
56	46.8	8.8	140884	9	AC139618	Homo sapi
57	46.8	8.8	252420	3	AE014841	Homo sapi
58	46.8	8.8	273275	3	AE014828	Plasmid
59	46.8	8.8	343050	3	PFA929353	Plasmid
60	46.6	8.8	82078	9	AC020379	Continuation
61	46.6	8.8	110000	2	AL732359_09	Continuation
62	46.6	8.8	110000	3	AC116984_3	Continuation
63	46.6	8.8	166420	2	AC012112	Continuation
64	46.6	8.8	252394	3	AE014833	Continuation
65	46.4	8.8	56080	8	AP003838	Continuation

[illegible]

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-1058B24"
/complement(2..45)
/rpt_family="THE1B"
413..435
/rpt_family="AT_rich"
442..918
/rpt_family="LIMAS"
1031..1232
/rpt_family="LIMB3"
complement(1386..1622)
/rpt_family="MIR"
1782..2075
/rpt_family="AluSq"
2157..2407
/rpt_family="LIPB1"
2408..2437
/rpt_family="(TG)n"
2437..2461
/rpt_family="(TA)n"
2462..2827
/rpt_family="LIPB1"
3113..3144
/rpt_family="(TG)n"
complement(5098..5729)
/rpt_family="LIMB8"
6062..6209
/rpt_family="LIM4"
6591..6668
/rpt_family="MER116"
complement(7056..7171)
/rpt_family="MIR"
7595..7933
/rpt_family="MLT1J2"
complement(8371..8498)
/rpt_family="LIMAS"
complement(8782..8958)
/rpt_family="MIR3"
10392..10495
/rpt_family="MIR"
complement(10523..10550)
/rpt_family="MIR"
10620..10688
/rpt_family="MIR"
11359..11423
/rpt_family="(TGGA)n"
11445..11611
/rpt_family="(TGGA)n"
complement(11724..11993)
/rpt_family="L3"
12093..12262
/rpt_family="LTR41"
complement(12751..12896)
/rpt_family="MIR"
12903..13006
/rpt_family="HAL1"
complement(13397..13512)
/rpt_family="L2"
13613..14101
/rpt_family="MLT2B1"
14267..14555
/rpt_family="AluSp"
complement(14845..15049)
/rpt_family="LIMCS"
complement(15054..15164)
/rpt_family="LIMCS"
16324..16627
/rpt_family="AluY"

repeat_region complement(17668..17733)
/rpt_family="L2"
repeat_region complement(18517..18621)
/rpt_family="MIR"
repeat_region 19516..20673
/rpt_family="LIME1"
repeat_region 20701..20841
/rpt_family="MIR3"
repeat_region 23592..24723
/rpt_family="LIMC2"
repeat_region complement(24737..24859)
/rpt_family="FLAM_C"
repeat_region 24904..24967
/rpt_family="L2"
repeat_region complement(27155..27449)
/rpt_family="AluSx"
repeat_region complement(27756..27864)
/rpt_family="L2"
repeat_region 28024..28283
/rpt_family="MIR"
repeat_region 28624..28864
/rpt_family="AluSq"
repeat_region 29251..29273

Query Match 97.8%; Score 517.6; DB 9; Length 117978;
Best Local Similarity 99.6%; Pred. No. 3e-104;
Matches 528; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CATGGTATGCTGTAGTCGGCTGATTTATATGCTGATTTATATGCTGATTTATGGTGATTTGCTTCT 60
DB 88017 CATGGTATGCTGTAGTCGGCTGATTTATATGCTGATTTATATGCTGATTTATGGTGATTTGCTTCT 88076
QY 61 TCTTTATACCTTTTATTTATTTCCCAAAATTTTCTTAAAGCAAAATATTTCTTGGCTAATCAAT 120
DB 88077 TCTTTATACCTTTTATTTATTTATTTCCCAAAATTTTCTTAAAGCAAAATATTTCTTGGCTAATCAAT 88136
QY 121 AAATTTATCAAAAGAAAAAAGAACTGAAAGCAACGCTTGAAAGAAAGAAAGTTAG-CCCTA 179
DB 88137 AAATTTATCAAAAGAAAAAAGAACTGAAAGCAACGCTTGAAAGAAAGAAAGTTAGCCCTA 88196
QY 180 TCGGGTATATTTTGGAGTTGTAAATACTACTGTCTCTTCTTAAAGTCCCACTCTCTG 239
DB 88197 TCGGGTATATTTTGGAGTTGTAAATACTACTGTCTCTTCTTAAAGTCCCACTCTCTG 88256
QY 240 TTTTCTTTGACAGCAAGAAAGAGCAAGTCAACCTCTCTTCCATGACAAACCAACCATGAC 299
DB 88257 TTTTCTTTGACAGCAAGAAAGAGCAAGTCAACCTCTCTTCCATGACAAACCAACCATGAC 88316
QY 300 CGGCTGTCTGCTCTCATCACAGGAAGGTGCAGAGAGCCCTTTTCTTCCAGGGTTG 359
DB 88317 CGGCTGTCTGCTCTCATCACAGGAAGGTGCAGAGAGCCCTTTTCTTCCAGGGTTG 88376
QY 360 TGGGCACTTGAATCTCTGCTTGGCTTGTAGCTAGAACTGCCCAAGTGGCACTC 419
DB 88377 TGGGCACTTGAATCTCTGCTTGGCTTGTAGCTAGAACTGCCCAAGTGGCACTC 88436
QY 420 AAGCTGTTTCGGGGACTTCCAGATAAAACCAACCAAGCTGTATTACTGTCTTTGACATCC 479
DB 88437 AAGCTGTTTCGGGGACTTCCAGATAAAACCAACCAAGCTGTATTACTGTCTTTGACATCC 88496
QY 480 TCTCTGGTCAGGGCCAAAGTCACTGTCTCTGCTTGGAGAGGGCTTCTGAG 529
DB 88497 TCTCTGGTCAGGGCCAAAGTCACTGTCTCTGCTTGGAGAGGGCTTCTGAG 88546

RESULT 2
AC023385/c
LOCUS 167118 bp DNA linear PRI 31-JUL-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-675P19, complete sequence.
ACCESSION AC023385
VERSION AC023385.10 GI:22024594
KEYWORDS HTG.
SOURCE Homo sapiens (human)
```

ORGANISM	Homo sapiens	Direct Submission	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Submitted (31-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
AUTHORS	1 (bases 1 to 167118)	On Jul 31, 2002 this sequence version replaced gi:21699264.	
TITLE	Homo sapiens chromosome 8, clone RP11-675P19	All repeats were identified using RepeatMasker:	
JOURNAL	Unpublished	Smit, A.F.A. & Green, P. (1996-1997)	
REFERENCE	2 (bases 1 to 167118)	http://ftp.genome.washington.edu/RM/RepeatMasker.html	
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beda, P., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castile, A., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Laroque, K., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivari, T.M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIER Web site: http://www-seq.wi.mit.edu Contact: sequence_submission@genome.wi.mit.edu ----- Project Information Center project name: L5118 Center clone name: 675_P19	
TITLE	Direct Submission	Location/Qualifiers	
JOURNAL	Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	1. 167118	
REFERENCE	3 (bases 1 to 167118)	/organism="Homo sapiens"	
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, A., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	/mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="8" /map="8" /clone="RP11-675P19" /clone_lib="RPC1-11 Human Male BAC" 968. 989 /rpt_family="AT_rich" 1357. 1540 /rpt_family="MERSA" complement(1546. 1591) /rpt_family="MLT1J2" complement(1592. 1655) /rpt_family="MLT1A1" complement(1675. 1844) /rpt_family="LIM4" complement(1860. 2187) /rpt_family="MER76" complement(2190. 2598) /rpt_family="LIM4" 4090. 4321 /rpt_family="MER1B" complement(4751. 5037) /rpt_family="AluJo" 6251. 6291 /rpt_family="GC_rich" 7341. 7504 /rpt_family="MERSA" 8072. 8096 /rpt_family="(A)n" 9408. 9791 /rpt_family="THE1C" 10075. 10393 /rpt_family="LIMC4a" 10963. 11107 /rpt_family="LIMC4a" complement(11124. 11360) /rpt_family="MIR" 11505. 12283 /rpt_family="LIMC4a" 15377. 15741 /rpt_family="MLT1A1" 17846. 17906 /rpt_family="(TATAA)n" complement(17909. 18160) /rpt_family="AluSx" complement(18412. 18708) /rpt_family="AluSx" complement(18755. 19749) /rpt_family="LIMC2" complement(19750. 20046) /rpt_family="AluSx" complement(20047. 21283) /rpt_family="LIMC2"	
TITLE	Direct Submission		
JOURNAL	Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	4 (bases 1 to 167118)		
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, A., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1
* 656: contig of 656 bp in length
* 657
* 756: gap of 100 bp
* 757
* 1450: contig of 694 bp in length
* 1451
* 1550: gap of 100 bp
* 1551
* 2232: contig of 682 bp in length
* 2233
* 2332: gap of 100 bp
* 2333
* 2983: contig of 651 bp in length
* 2984
* 3083: gap of 100 bp
* 3084
* 3725: contig of 642 bp in length
* 3726
* 3825: gap of 100 bp
* 3826
* 4455: contig of 630 bp in length
* 4456
* 4555: gap of 100 bp
* 4556
* 5212: contig of 657 bp in length
* 5213
* 5312: gap of 100 bp
* 5313
* 5978: contig of 666 bp in length
* 5979
* 6078: gap of 100 bp
* 6079
* 6758: contig of 680 bp in length
* 6759
* 6858: gap of 100 bp
* 6859
* 7534: contig of 676 bp in length
* 7535
* 7634: gap of 100 bp
* 7635
* 8307: contig of 673 bp in length
* 8308
* 8407: gap of 100 bp
* 8408
* 9064: contig of 657 bp in length
* 9065
* 9164: gap of 100 bp
* 9165
* 9843: contig of 679 bp in length
* 9844
* 9943: gap of 100 bp
* 9944
* 10631: contig of 688 bp in length
* 10632
* 10731: gap of 100 bp
* 10732
* 11419: contig of 688 bp in length
* 11420
* 11519: gap of 100 bp
* 11520
* 12185: contig of 666 bp in length
* 12186
* 12285: gap of 100 bp
* 12286
* 12931: contig of 646 bp in length
* 12932
* 13031: gap of 100 bp
* 13032
* 13700: contig of 669 bp in length
* 13701
* 13800: gap of 100 bp
* 13801
* 14482: contig of 682 bp in length
* 14483
* 14582: gap of 100 bp
* 14583
* 15261: contig of 679 bp in length
* 15262
* 15361: gap of 100 bp
* 15362
* 16042: contig of 681 bp in length
* 16043
* 16142: gap of 100 bp
* 16143
* 16823: contig of 681 bp in length
* 16824
* 16923: gap of 100 bp
* 16924
* 17615: contig of 692 bp in length
* 17616
* 17715: gap of 100 bp
* 17716
* 18373: contig of 658 bp in length
* 18374
* 18473: gap of 100 bp
* 18474
* 19160: contig of 687 bp in length
* 19161
* 19260: gap of 100 bp
* 19261
* 19932: contig of 672 bp in length
* 19933
* 20032: gap of 100 bp
* 20033
* 20689: contig of 657 bp in length
* 20690
* 20789: gap of 100 bp
* 20790
* 21446: contig of 657 bp in length
* 21447
* 21546: gap of 100 bp
* 21547
* 22186: contig of 640 bp in length
* 22187
* 2286: gap of 100 bp
* 2287
* 22973: contig of 687 bp in length
* 22974
* 23073: gap of 100 bp
* 23074
* 23762: contig of 689 bp in length
* 23763
* 23862: gap of 100 bp
* 23863
* 24538: contig of 676 bp in length
* 24539
* 24638: gap of 100 bp
* 24639
* 25309: contig of 671 bp in length
* 25310
* 25409: gap of 100 bp

* 25410
* 26102
* 26201: contig of 692 bp in length
* 26202: gap of 100 bp
* 26893: contig of 692 bp in length
* 26894
* 26993: gap of 100 bp
* 26994
* 27685: contig of 692 bp in length
* 27686
* 27785: gap of 100 bp
* 27786
* 28456: contig of 671 bp in length
* 28457
* 28556: gap of 100 bp
* 28557
* 29193: contig of 637 bp in length
* 29194
* 29293: gap of 100 bp
* 29294
* 29970: contig of 677 bp in length
* 29971
* 30070: gap of 100 bp
* 30071
* 30757: contig of 687 bp in length
* 30758
* 30857: gap of 100 bp
* 30858
* 31520: contig of 663 bp in length
* 31521
* 31620: gap of 100 bp
* 31621
* 32313: contig of 693 bp in length
* 32314
* 32413: gap of 100 bp
* 32414
* 33074: contig of 661 bp in length
* 33075
* 33174: gap of 100 bp
* 33175
* 33880: contig of 706 bp in length
* 33881
* 33980: gap of 100 bp
* 33981
* 34667: contig of 687 bp in length
* 34668
* 34767: gap of 100 bp
* 34768
* 35390: contig of 623 bp in length
* 35391
* 35490: gap of 100 bp
* 35491
* 36141: contig of 651 bp in length
* 36142
* 36241: gap of 100 bp
* 36242
* 36873: contig of 632 bp in length
* 36874
* 36973: gap of 100 bp
* 36974
* 37642: contig of 669 bp in length
* 37643
* 37742: gap of 100 bp
* 37743
* 38426: contig of 684 bp in length
* 38427
* 38526: gap of 100 bp
* 38527
* 39190: contig of 664 bp in length
* 39191
* 39290: gap of 100 bp
* 39291
* 39592: contig of 662 bp in length
* 39593
* 40052: gap of 100 bp
* 40053
* 40712: contig of 660 bp in length
* 40713
* 40812: gap of 100 bp
* 40813
* 41470: contig of 658 bp in length
* 41471
* 41570: gap of 100 bp
* 41571
* 42240: contig of 670 bp in length
* 42241
* 42340: gap of 100 bp
* 42341
* 43014: contig of 674 bp in length
* 43015
* 43114: gap of 100 bp
* 43115
* 43762: contig of 648 bp in length
* 43763
* 43862: gap of 100 bp
* 43863
* 44496: contig of 634 bp in length
* 44497
* 44596: gap of 100 bp
* 44597
* 45274: contig of 678 bp in length
* 45275
* 45374: gap of 100 bp
* 45375
* 46071: contig of 697 bp in length
* 46072
* 46171: gap of 100 bp
* 46172
* 46850: contig of 679 bp in length
* 46851
* 46950: gap of 100 bp
* 46951
* 47625: contig of 675 bp in length
* 47626
* 47725: gap of 100 bp
* 47726
* 48419: contig of 694 bp in length
* 48420
* 48519: gap of 100 bp
* 48520
* 49175: contig of 656 bp in length
* 49176
* 49275: gap of 100 bp
* 49276
* 49967: contig of 692 bp in length
* 49968
* 50067: gap of 100 bp
* 50068
* 50752: contig of 685 bp in length
* 50753
* 50852: gap of 100 bp
* 50853
* 51504: contig of 652 bp in length
* 51505
* 51604: gap of 100 bp
* 51605
* 52254: contig of 650 bp in length
* 52255
* 52354: gap of 100 bp
* 52355
* 53006: contig of 652 bp in length
* 53007
* 53106: gap of 100 bp
* 53107
* 53786: contig of 680 bp in length

Qy	3	TGGTGATGCTAGTCGTGGCGCATTTATACTGATTATGGGTGAATTTTGCCTCCCTTC	62
Db	169225	TGTTTATTAATTTTTTTTTTATTTTTTATTTTTTATTTATTTATTTATTTATTTA	169284
Qy	63	TTTATACCTTTTATTTATTTCCCAAAATTTTTCTTTAACGAAATATTTCTTTGCTAATCAATAA	122
Db	169285	TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTAAAAGTCACAGAGAAAAAGAAA	169344
Qy	123	ATTATCAAAAGAAAAAAAACCTGAAAGCAACGCTTTGAAAAAGGAAAGTAGTACCCTATCG	182
Db	169345	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGGAAAAAAAAAAAAAAAAAAAAATGA	169404
Qy	183	GGTATATTTTGGAGAGTTGTAATAA	205
Db	169405	AAAAAGATTTTAAATAATCCAATAA	169427
RESULT 14			
DDIDDP2/c			
LOCUS	DDIDDP2	5850 bp	DNA linear INV 27-APR-1993
DEFINITION	Dictyostelium discoideum plasmid Ddp2 trans-acting factor gene,		
	complete cds.		
ACCESSION	M55298		
VERSION	M55298.1 GI:167727		
KEYWORDS	trans-acting factor;		
SOURCE	Dictyostelium discoideum		
ORGANISM	Dictyostelium discoideum		
REFERENCE	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.		
AUTHORS	Leiting,B., Lindner,I.J. and Noegel,A.A.		
TITLE	The extrachromosomal replication of Dictyostelium plasmid Ddp2 requires a cis-acting element and a plasmid-encoded trans-acting factor		
JOURNAL	Mol. Cell. Biol. 10 (7), 3727-3736 (1990)		
MEDLINE	90287164		
PUBMED	2192261		
COMMENT	Original		source text: Dictyostelium discoideum (strain WS380B) DNA.
FEATURES	Location/Qualifiers		
source	1..5850		
	/organism="Dictyostelium discoideum"		
	/mol_type="genomic DNA"		
	/strains="WS380B"		
	/db_xref="taxon:44689"		
	1795..1829		
	/note="putative VECTOR sequence Bacteriophage lambda (J02459); putative"		
	2372..5041		
	/codon_start=1		
	/product="trans-acting factor"		
	/protein_id="AAA33191.1"		
	/db_xref="GI:167728"		
	/translation="MDELI SWDRPFKPVVILLEEFGCKRNDVRLSDVDYDILSGIYSP RTFVLKEVFRAVAVSYDESEIDLRLGSVPFGTSLSYVPGIFSLKDPLLISKTSGK IRVSDVOAILPHFPSRISSQVFRKDIIPGYTFEKSLSSEYKISGRAGVSWFN LVSKISTYKNHPFAENPTKYHDFISMLSVHGIIIVDSQNEEDNNYSAMYSLPFF DLKSDIPGAVQSRTSRKSGSLKDVIRNNISVDSSESDEVMSVSNELKVLEAVNKG SVQQSSRQPSTSRKSGSLKDVIRNNISVDSSESDEVMSVSNELKVLEAVNKG EIGEVVKVLTLQDKNYSGLLENIEHNKSERVITLSSFPFIASKINYDEVKFS ELSDLVESAKELFEKNTNILLIPTNFKEGFELWPIVINGIASTSVFSPNNYSSG SPANEGALKIHICISIGNNGFLSIKSTIFTDPFKISTKDLIPMSRKMLDLEQQFRX LRDANNNSKKSKVQSDSDIGTDIEDTKLISFVHFINDNLKLKSKEDGLMLVDYFP TSLFMRYNPNSIDNKGMFMCHEISEIKFSQCKNHSIDNLVLSFTPNNIKINSOINE NELKKYSLAWSDFRNVPKTPKPIPSBFKFTITFTNNSYNARVAFDDISGIS ITNVKNIHAKQRPAFEIYETLLGSTRIIRAFPCAPCLIQINNKFATDKLIDDOSVNHI QIASLEIKNLSPLPIDIKVRGSTGVTIKGGTAPIINSEDTFSSICLDIRFSASLI SKTSLSQTPAPBERYNKETNLIKLVQDCDELFTFLNNYKIANKLKSTIERLYNNF MGLDE complement(5073..5113) /note="putative VECTOR sequence Cloning vector pEMBL 8 minus (pEMBLam) (X04596); putative"		
misc_feature			
CDS			
BASE COUNT	2300 a	646 c	706 g 2198 t
ORIGIN	Plasmid Ddp2.		


```
/notes="palindromic sequence tggcatgaca"
2378. .5041
/genes="rep"
CDS
2378. .5041
/genes="rep"
/notes="essential for plasmid replication"
/codon_start=1
/protein_id="CAA35843.1"
/db_xref="GI:7308"
/db_xref="SPTREMBL:Q23913"
/translation="MDELISWDRPFKFFVILLSEPKGCKRNDVRLSDVYDILSGIYSP
RTFVLKEVRAVAVSYDESEIDLFRUGSVFPFGTSLSYIPGIFSLKDFLLISKTSGK
IRVSDVQALLIFDHFSDRSIDKQVFRKDIIPGRTFPEKSISEYKISDGRAAGVSWFN
LVSKISTYCKNHPFAENPTYKHVDFISMLSLVHGIIIVDSQNEDENNVAMTSLNPFV
DLEKSDIPCAVQSRVTNTRGRSNNLNPTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
SVQEOSRQPKTSRSGSLKDVIRNNISVDSSESDDVMSVSNRLKCVLLRAVNVKG
EIGLEVVKVLQDKQNYATGLLENIENHNSERVITLSSFFPEIASKINDYDEVKFS
ELSIDVESAKRLTFEKNITNLIPTNNKEGFEFLWPIVNGIASTSVFSPNNYSSG
SPANVESALKLHLCSIGNINGFUSIRITFDTKSITKDLIPMSKRMMLDLEQGFPRK
LTDANNNSKSKVQSDISGIDETDKLISFVHEFINDNLYKLKSKEDGMLVDVFP
TSTLFRNPNNSIDNKVGMFHCBSISKFSQCKNHSIDNVLSTPNNIKNISODNE
NELKKYSIMVSDFRNVKVPKFIPESEKRTIITFTNNSYNANRVAFDDISGSI
ITNVKNIHAKGORNFIETLGLSTRIRAFPCAPCLIQINNFKPATDKLIDDQSVNH
QTASLEIKNLSYPLDIDKVRGTVGTIGKETAPIIINSEETFFSICLDIRFSALI
SKTKLSQLPFTAPDERYNKTNILKVLQDCDELTRFLNNYKIANKLSTIENLYNNF
MGLEDEDEDEDEDEDEDEDEDEGY"
BASE COUNT 2296 a 652 c 709 g 2195 t
ORIGIN

Query Match 9.3%; Score 49.2; DB 3; Length 5852;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 12 TGTAGTCGGCTGATTATATCTGATTATGCGGATTTGCTTCTTTTACTT 71
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5767 TTTAATATTGTTATTTGTTATTTTATATATGTTATTTGTTGTTTACTT 5708
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 72 TTATTTATCCCAAAATTTCTTAAAGCAAAATTTCTTCTTAATCAATAATATCAAA 131
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5707 ATATTTCTATTTTATTTATATAAATTAATAATTAATAATAAATAAATAA 5648
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 132 AGAAAAAACTGAAGCAACGCTGAAAAAGGAAA 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5647 AAAAAAAATTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 5610
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 16
A21622/c
LOCUS
DEFINITION Dictoselium plasmid Ddp2 Rep gene.
ACCESSION A21622
VERSION A21622.1 GI:583618
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM Dictoselium spp.
REFERENCE 1 (bases 1 to 5852)
AUTHORS
TITLE IMPROVED PLASMID VECTORS FOR CELLULAR SLIME MOULDS OF THE GENUS
DICTYOSTELIUM
JOURNAL Patent: WO 9106644-A 14 16-MAY-1991;
FEATURES
source
1. .5852
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
2378. .5041
/genes="REP gene from a recombinant Ddp-2 like plasmid in
Dictyostelium spp"
CDS
2378. .5041
/genes="REP gene from a recombinant Ddp-2 like plasmid in
Dictyostelium spp"
/codon_start=1

/transl_table=11
/protein_id="CAA01565.1"
/db_xref="GI:583619"
/translation="MDELISWDRPFKFFVILLSEPKGCKRNDVRLSDVYDILSGIYSP
RTFVLKEVRAVAVSYDESEIDLFRUGSVFPFGTSLSYIPGIFSLKDFLLISKTSGK
IRVSDVQALLIFDHFSDRSIDKQVFRKDIIPGRTFPEKSISEYKISDGRAAGVSWFN
LVSKISTYCKNHPFAENPTYKHVDFISMLSLVHGIIIVDSQNEDENNVAMTSLNPFV
DLEKSDIPCAVQSRVTNTRGRSNNLNPTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
SVQEOSRQPKTSRSGSLKDVIRNNISVDSSESDDVMSVSNRLKCVLLRAVNVKG
EIGLEVVKVLQDKQNYATGLLENIENHNSERVITLSSFFPEIASKINDYDEVKFS
ELSIDVESAKRLTFEKNITNLIPTNNKEGFEFLWPIVNGIASTSVFSPNNYSSG
SPANVESALKLHLCSIGNINGFUSIRITFDTKSITKDLIPMSKRMMLDLEQGFPRK
LTDANNNSKSKVQSDISGIDETDKLISFVHEFINDNLYKLKSKEDGMLVDVFP
TSTLFRNPNNSIDNKVGMFHCBSISKFSQCKNHSIDNVLSTPNNIKNISODNE
NELKKYSIMVSDFRNVKVPKFIPESEKRTIITFTNNSYNANRVAFDDISGSI
ITNVKNIHAKGORNFIETLGLSTRIRAFPCAPCLIQINNFKPATDKLIDDQSVNH
QTASLEIKNLSYPLDIDKVRGTVGTIGKETAPIIINSEETFFSICLDIRFSALI
SKTKLSQLPFTAPDERYNKTNILKVLQDCDELTRFLNNYKIANKLSTIENLYNNF
MGLEDEDEDEDEDEDEDEDEDEGY"
BASE COUNT 2296 a 652 c 709 g 2195 t
ORIGIN

Query Match 9.3%; Score 49.2; DB 6; Length 5852;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 12 TGTAGTCGGCTGATTATATCTGATTATGCGGATTTGCTTCTTTTACTT 71
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5767 TTTAATATTGTTATTTGTTATTTTATATATGTTATTTGTTGTTTACTT 5708
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 72 TTATTTATCCCAAAATTTCTTAAAGCAAAATTTCTTCTTAATCAATAATATCAAA 131
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5707 ATATTTCTATTTTATTTATATAAATTAATAATTAATAATAAATAAATAA 5648
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 132 AGAAAAAACTGAAGCAACGCTGAAAAAGGAAA 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5647 AAAAAAAATTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 5610
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 17
AC116977/c
LOCUS
DEFINITION Dictyostelium discoideum chromosome 2 map 5515173-5817331 strain
AX4, complete sequence.
AC116977
VERSION AC116977.2 GI:28828573
KEYWORDS HTG.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 302156)
AUTHORS Gloeckner, G., Eichinger, L., Szafranski, K., Pachebat, J., Dear, P.,
Lehmann, R., Baumgart, C., Parra, G., April, J. F., Guigo, R., Kumpf, K.,
Tunggal, B., Cox, E., Quail, M. A., Platzer, M., Rosenthal, A. and
Noegel, A. A.
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum
JOURNAL Nature 418 (6893), 79-85 (2002)
MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
REFERENCE 2 (bases 1 to 302156)
AUTHORS Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 302156)
AUTHORS Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 302156)
AUTHORS Baumgart, C.
TITLE Direct Submission
```



```

/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-624D11"
/clone_lib="RPC1-11 Human Male BAC"
997. .917
/rpt_family="AT_rich"
1150. .1169
/rpt_family="(T)n"
complement(1820. .1987)
/rpt_family="MIR"
complement(2145. .2206)
/rpt_family="MIR"
3207. .3495
/rpt_family="L2"
3995. .4080
/rpt_family="MIR3"
4157. .4648
/rpt_family="MLT1E"
4894. .4958
/rpt_family="MIR3"
4962. .5018
/rpt_family="MER5A"
5019. .5783
/rpt_family="L1P4"
5784. .5903
/rpt_family="MER5A"
complement(5956. .6069)
/rpt_family="MIR"
complement(6952. .7163)
/rpt_family="MIR"
complement(7264. .7478)
/rpt_family="MLT1-int"
7479. .7761
/rpt_family="AluSc"
complement(7762. .8046)
/rpt_family="MLT1-int"
complement(8103. .8751)
/rpt_family="MLT1-int"
complement(8770. .9068)
/rpt_family="MLT2C1"
9069. .9368
/rpt_family="AluSx"
complement(9369. .9480)
/rpt_family="MLT2C1"
9634. .9925
/rpt_family="AluJo"
9997. .10128
/rpt_family="AluSp/q"
complement(11650. .11717)
/rpt_family="LTR33"
11726. .12087
/rpt_family="MLT1A2"
12577. .12598
/rpt_family="(CAAA)n"
12688. .12777
/rpt_family="Charlie8"
complement(12928. .13032)
/rpt_family="MIR"
13088. .13466
/rpt_family="L1M4"
15240. .15264
/rpt_family="AT_rich"
complement(15331. .15556)
/rpt_family="MIR"
complement(15708. .15842)
/rpt_family="L2"
18201. .18387
/rpt_family="MER117"
complement(18507. .18727)
/rpt_family="MIR"
complement(19164. .19252)
/rpt_family="L2"

repeat_region complement(21582. .21799)
/rpt_family="MIR"
repeat_region complement(22125. .22227)
/rpt_family="L1M4"
repeat_region 22264. .22287
/rpt_family="AT_rich"
repeat_region complement(22347. .22466)
/rpt_family="LTR16C"
repeat_region complement(22698. .22946)
/rpt_family="L1M4"
repeat_region 22947. .23343
/rpt_family="MLT1J1"
repeat_region 24245. .24502
/rpt_family="MIR"
repeat_region 25311. .25454
/rpt_family="L3"
repeat_region complement(25481. .25513)
/rpt_family="MER5A"
repeat_region 25514. .25533
/rpt_family="(CAAA)n"
repeat_region complement(25534. .25598)
/rpt_family="MER5A"
repeat_region complement(25599. .26466)

Query Match 9.1%; Score 48.2; DB 9; Length 155985;
Best Local Similarity 52.3%; Pred. No. 1.8;
Matches 127; Conservative 1; Mismatches 114; Indels 1; Gaps 1;

QY 26 ATTTATATGCTGATTATGGGTGATTTTGGTCTCTCTTTATATCTTTATTTATTTATTTCCCAA 85
DB 57319 ATCACTATACCTGATTATGAATGATTTATTTCTCTTTGTTATATGATTTCTCTATTTCCAAA 57260
QY 86 ATTTTCTTAAAGCAATATTT-CTTGTCTATCAATAAATTTATCAAAAGAAAAAAACT 144
DB 57259 CATTATGCATTAACATATTTACTTTTATTAATAATGATCATAGAAAAAAGGGA 57200
QY 145 GAAAGCAACGCTTGAAAAAAGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAAA 204
DB 57199 AACGTCACATCATGATTTATGACACAAAATTTCTGTATTTAGTTGCTACTCTCCAC 57140
QY 205 ATACTACTGTCTCTCTAAGTCCCACTCTCTGTTTCTTTGAGCAGGAAGAAG 264
DB 57139 AAAACAGAGCTTCATTTTACAATGACAGACCCCTTTATCTCGATAGGAAGGGGAAA 57080
QY 265 CAG 267
DB 57079 AAG 57077

RESULT 25
AC068749
LOCUS AC068749 181996 bp DNA linear HTG 04-MAY-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-802012 map 11, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION AC068749
VERSION AC068749.2 GI:8389606
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 181996)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, P.,
Boguslavskiy, L., Boukhgelter, B., Brown, A., Burkett, G.,
Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

```

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
 Meidrich,J., Meneus,L., Mhova,T., Miranda,C., Mienga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Submitted (08-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 9, 2000 this sequence version replaced gi:7717156.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L9955

Center clone name: 802_O12

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 171741 bases at least Q40

Consensus quality: 176943 bases at least Q30

Consensus quality: 179052 bases at least Q20

Insert size: 182000; agarose-fp

Quality coverage: 4.5 in Q20 bases; sum-of-contigs

Quality coverage: 4.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1109: contig of 1109 bp in length
 * 1110 1209: gap of 100 bp
 * 1210 2970: contig of 1761 bp in length
 * 2971 3070: gap of 100 bp
 * 3071 8242: contig of 5172 bp in length
 * 8243 8342: gap of 100 bp
 * 8343 14456: contig of 6114 bp in length
 * 14457 14556: gap of 100 bp
 * 14557 19217: contig of 4661 bp in length
 * 19218 19317: gap of 100 bp
 * 19318 24451: contig of 5134 bp in length
 * 24452 24551: gap of 100 bp
 * 24552 30373: contig of 5822 bp in length
 * 30374 30473: gap of 100 bp
 * 30474 38117: contig of 7644 bp in length
 * 38118 38217: gap of 100 bp
 * 38218 45565: contig of 7348 bp in length
 * 45566 45666: gap of 100 bp
 * 45666 57699: contig of 12034 bp in length
 * 57700 57700: gap of 100 bp
 * 57800 70834: contig of 13035 bp in length
 * 70835 70835: gap of 100 bp
 * 70935 84875: contig of 13941 bp in length
 * 84876 84975: gap of 100 bp

* 84976 98366: contig of 13391 bp in length
 * 98367 98466: gap of 100 bp
 * 98467 117242: contig of 18776 bp in length
 * 117243 117342: gap of 100 bp
 * 117343 147820: contig of 30478 bp in length
 * 147821 147920: gap of 100 bp
 * 147921 181996: contig of 34076 bp in length.

FEATURES

source

Location/Qualifiers

1..181996

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="RP11-802Q12"

/clone_lib="RPC1-11 Human Male BAC"

misc_feature

1..1109

/note="assembly_fragment"

misc_feature

1210..2970

/note="assembly_fragment"

misc_feature

3071..8242

/note="assembly_fragment"

misc_feature

8343..14456

/note="assembly_fragment"

misc_feature

14557..19217

/note="assembly_fragment"

misc_feature

clone_end:17

vector_side:right

misc_feature

19318..24451

/note="assembly_fragment"

misc_feature

24552..30373

/note="assembly_fragment"

misc_feature

30474..38117

/note="assembly_fragment"

misc_feature

38218..45565

/note="assembly_fragment"

misc_feature

45666..57699

/note="assembly_fragment"

misc_feature

57800..70834

/note="assembly_fragment"

misc_feature

70935..84875

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

clone_end:SP6

vector_side:left

misc_feature

84976..98366

/note="assembly_fragment"

misc_feature

98467..117242

/note="assembly_fragment"

misc_feature

117343..147820

/note="assembly_fragment"

misc_feature

147921..181996

/note="assembly_fragment"

misc_feature

55224 a 33965 c 35178 g 56119 t 1510 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 9.1%; Score 48.2; DB 2; Length 181996;

Matches 127; Conservative 1; Mismatches 114; Indels 1; Gaps 1;

QY 26 ATTATATGCTGATTTATGGGTGATTTGCTTCTTTATACCTTTATTTATTTTCCCAA 85

DB 9746 ATCAGTATACCTGATTTATGAATGATTTTATTTCTTTTATGATTTCTCTATTTCCAAA 9805

QY 86 ATTTTCTTAAGCAATATTTT-CTTTGCTTAATCAATAATTTATCAAAAGAAAAAACT 144

DB 9806 CATTATGCAATAACATATTTACTTTTATTAATAATGATATCATAGAAAAAGGAA 9865

QY 145 GAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGAAGTTGTAAA 204

DB 9866 AAGGCTACTACATTGAGTATTATGACACAAAAATTTCTGTTATTTAGATTGCTACTCTCCAC 9925

QY 205 ATACTACGTGTTCTTCTTAAGTCCCACTCTCTGTTTTCTTTGAGGAGGAAGAAG 264


```

Db      9926 AAAACAGAGCTTCATTTTACATGACAGACCCCTTTATCTGAGATAGAGAAAGGGGAAA 9985
QY      265 CAG 267
        ||
Db      9986 AAG 9988

RESULT 26
AL356371
LOCUS   AL356371          256221 bp    DNA    linear    HTG 20-OCT-2001
DEFINITION Homo sapiens chromosome 11 clone RP1-135I9, *** SEQUENCING IN
PROGRESS ***, 22 unordered pieces.
ACCESSION AL356371
VERSION   AL356371.18 GI:14575185
KEYWORDS  HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
          McClay, K.
          Direct Submission
          Submitted (16-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
          CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
          requests: clonerequest@sanger.ac.uk
          On Jun 28, 2001 this sequence version replaced gi:14529810.
          ----- Genome Center
          Center: Sanger Centre
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: humquery@sanger.ac.uk
          ----- Project Information
          Center project name: d113519
          ----- Summary Statistics
          Assembly program: XGAP4; version 4.5
          Sequencing vector: M13; M77815; 1% of reads
          Sequencing vector: plasmid; L08752; 98% of reads
          Chemistry: Dye-terminator ET-amersham; 2% of reads Chemistry:
          Dye-terminator Big Dye; 97% of reads
          Consensus quality: 250889 bases at least Q40
          Consensus quality: 252160 bases at least Q30
          Consensus quality: 252999 bases at least Q20
          Insert size: 254121; sum-of-contigs
          Insert size: 98381; 1.2% error; agarose-fp
          Quality coverage: 9.35x in Q20 bases; sum-of-contigs Quality
          coverage: 25.82x in Q20 bases; agarose-fp
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 22 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *
          1 2923: contig of 2923 bp in length
          * 2924: gap of 100 bp
          * 3024: contig of 3307 bp in length
          * 6331: gap of 100 bp
          * 6431: contig of 7855 bp in length
          * 14286: contig of 100 bp
          * 14386: gap of 100 bp
          * 26365: contig of 11980 bp in length
          * 26366: gap of 100 bp
          * 26466: contig of 11549 bp in length
          * 38015: gap of 100 bp
          * 41035: contig of 2920 bp in length
          * 41135: gap of 100 bp
          * 52552: contig of 11417 bp in length
          * 52651: gap of 100 bp
          * 52652: contig of 4140 bp in length
          * 56792: contig of 100 bp
          * 56892: gap of 100 bp
          * 65134: contig of 8242 bp in length
          * 65233: gap of 100 bp

          65234: contig of 2011 bp in length
          67344: gap of 100 bp
          69441: contig of 2097 bp in length
          69541: gap of 100 bp
          91678: contig of 22137 bp in length
          91779: gap of 100 bp
          97058: contig of 5280 bp in length
          97158: gap of 100 bp
          99624: contig of 2466 bp in length
          99724: gap of 100 bp
          106782: contig of 7058 bp in length
          106882: gap of 100 bp
          112388: contig of 5506 bp in length
          112488: gap of 100 bp
          116274: contig of 3786 bp in length
          116374: gap of 100 bp
          118939: contig of 2565 bp in length
          119040: gap of 100 bp
          129647: contig of 10608 bp in length
          129747: gap of 100 bp
          136936: contig of 7189 bp in length
          137036: gap of 100 bp
          248507: contig of 11471 bp in length
          248607: gap of 100 bp
          256221: contig of 7614 bp in length.

FEATURES             Location/Qualifiers
     source            1..256221
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="RP1-135I9"
                     /clone_lib="RPC1-1"
     misc_feature       1..2923
                     /note="assembly_fragment:00208"
     misc_feature       3024..6330
                     /note="assembly_fragment:00427"
     misc_feature       6431..14285
                     /note="assembly_fragment:00726"
     misc_feature       14386..26365
                     /note="assembly_fragment:00955.0"
     misc_feature       26466..38014
                     /note="assembly_fragment:01016"
     misc_feature       38115..41034
                     /note="assembly_fragment:01599.0"
     misc_feature       41135..52551
                     /note="assembly_fragment:02394"
     misc_feature       52652..56791
                     /note="assembly_fragment:02491.0"
     misc_feature       56892..65133
                     /note="assembly_fragment:02823"
     misc_feature       65234..67244
                     /note="assembly_fragment:03451"
     misc_feature       67345..69441
                     /note="assembly_fragment:03517"
     misc_feature       69542..91678
                     /note="assembly_fragment:03857"
     misc_feature       91779..97058
                     /note="assembly_fragment:04538"
     misc_feature       97159..99624
                     /note="assembly_fragment:04662"
     misc_feature       99725..106782
                     /note="assembly_fragment:04852"
     misc_feature       106883..112388
                     /note="assembly_fragment:04898"
     misc_feature       112489..116274
                     /note="assembly_fragment:05674.0"
     misc_feature       116375..118939
                     /note="assembly_fragment:06302"
     misc_feature       119040..129647
                     /note="assembly_fragment:06497"
     misc_feature       129748..136936
                     /note="assembly_fragment:06540"

```



```
/note="assembly_fragment
clone end:T7
vector_side:left"
85817..119636
/note="assembly_fragment"
119737..152336
/note="assembly_fragment"
BASE COUNT 51773 a 29398 c 27604 g 42706 t 855 others
ORIGIN
Query Match          9.0%; Score 47.8; DB 2; Length 152336;
Best Local Similarity 63.8%; Pred. No. 2.2;
Matches 90; Conservative 0; Mismatches 47; Indels 4; Gaps 1;
QY 26 ATTATATGCTGATTATGCTGATTTTGTCTCTTTATATCTTTTATTATTCCTCA 85
DB 106738 ATTACATCTGTAATATGATGATTTTCTTCTCTTTAGATTTTGGATTTCTCA 106797
QY 86 ATTTTCTCT---TAAGCAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAA 141
DB 106798 AATTTTCTATCTGAGTATGATTTATTTGTATAATCAGGAGTAGCAATTCAAAAAAGA 106857
QY 142 ACTGAAGCAACGCTTGAAAA 162
DB 106858 AAGAAAAATCTGCTTAATAA 106878
RESULT 30
AC079354 LOCUS 162554 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-1145F21 from 2, complete sequence.
ACCESSION AC079354
VERSION AC079354.4 GI:18098550
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 162554)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 162554)
AUTHORS Armstrong, J., Cotton, M., Hawkins, M. and Dignam, G.
TITLE The sequence of Homo sapiens BAC clone RP11-1145F21
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 162554)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 162554)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 9, 2002 this sequence version replaced gi:14550326.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH1145F21
-----
```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frenken, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-107N15, 2000 bp overlap; the clone sequenced to the right is RP11-686O6. Actual start of this clone is at base position 1 of RP11-1145F21; actual end is at base position 162554 of RP11-1145F21.

Data from AC069148 and AC064836 was used to finish this clone, AC079354.

FEATURES	Location/Qualifiers
source	1..162554
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-1145F21"
	/clone_lib="RPC1-11"
repeat_region	329..433
	/rpt_family="MIR"
repeat_region	1453..1493
	/rpt_family="(CA)n"
repeat_region	1668..1775
	/rpt_family="MIR"
repeat_region	1844..1916
	/rpt_family="LI"
misc_feature	2232..2402
	/note="similar to EST BB614631 (NID:gl5395798)"
repeat_region	2583..2603
	/rpt_family="AT-rich"
repeat_region	3403..3542
	/rpt_family="LI"
repeat_region	3568..3595
	/rpt_family="AT-rich"
repeat_region	3607..4123
	/rpt_family="ERV1"
repeat_region	4423..4588
	/rpt_family="(GGAA)n"
repeat_region	4608..5023
	/rpt_family="MaLR"
misc_feature	5049..5224
	/note="similar to EST BB614631 (NID:gl5395798)"
repeat_region	5692..5991
	/rpt_family="Alu"
repeat_region	5973..6004
	/rpt_family="AT-rich"
repeat_region	6261..6476
	/rpt_family="MIR"
repeat_region	7016..7295


```
FEATURES
source
    true right end of clone RP11-317C20 is at 56073 in this sequence.
    Location/Qualifiers
    1. .182012
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="9"
    /map="q31.1-31.3"
    /clone="RP11-413C10"
    /clone_lib="RPCI-11.2"
    2456. .2511
    /notes="28 copies 2 mer aa 71% conserved"
    repeat_region
    2622. .2787
    /notes="MIR repeat: matches 20. .196 of consensus"
    repeat_region
    2907. .3078
    /notes="L2 repeat: matches 2566. .2749 of consensus"
    repeat_region
    3110. .3345
    /notes="MIR repeat: matches 21. .237 of consensus"
    repeat_region
    3927. .4102
    /notes="88 copies 2 mer tt 57% conserved"
    repeat_region
    4108. .4406
    /notes="AluX repeat: matches 1. .303 of consensus"
    repeat_region
    6277. .6330
    /notes="27 copies 2 mer ta 72% conserved"
    repeat_region
    7275. .7467
    /notes="MIR repeat: matches 10. .219 of consensus"
    repeat_region
    7879. .7995
    /notes="LIPAS repeat: matches 6029. .6145 of consensus"
    repeat_region
    9391. .9426
    /notes="18 copies 2 mer tc 94% conserved"
    repeat_region
    9525. .9811
    /notes="MER47A repeat: matches 46. .333 of consensus"
    repeat_region
    10892. .10959
    /notes="L1M1 repeat: matches 718. .785 of consensus"
    repeat_region
    10998. .11075
    /notes="3 copies 26 mer 92% conserved"
    repeat_region
    11111. .11148
    /notes="19 copies 2 mer ac 92% conserved"
    repeat_region
    11507. .12390
    /notes="LIPAS repeat: matches 5254. .6143 of consensus"
    repeat_region
    15656. .16353
    /notes="L1M3 repeat: matches 5397. .6149 of consensus"
    repeat_region
    16387. .16491
    /notes="FLAM A repeat: matches 22. .127 of consensus"
    repeat_region
    16557. .16682
    /notes="L1M3 repeat: matches 5189. .5329 of consensus"
    repeat_region
    17439. .17610
    /notes="MIR repeat: matches 49. .241 of consensus"
    repeat_region
    19584. .20420
    /notes="LIPAS repeat: matches 5308. .6143 of consensus"
    repeat_region
    20813. .21163
    /notes="THB1A repeat: matches 1. .354 of consensus"
    repeat_region
    21164. .22730
    /notes="THB1A-internal repeat: matches 1. .1578 of consensus"
    repeat_region
    22735. .23078
    /notes="THB1A repeat: matches 3. .354 of consensus"
    repeat_region
    23719. .24027
    /notes="AluSq repeat: matches 1. .311 of consensus"
    repeat_region
    25619. .25765
    /notes="MER5A repeat: matches 15. .187 of consensus"
    repeat_region
    25784. .25850
    /notes="MER5B repeat: matches 103. .173 of consensus"
    repeat_region
    26130. .26262
    /notes="MER5A repeat: matches 7. .141 of consensus"
    repeat_region
    26884. .27199
    /notes="AluSq repeat: matches 1. .312 of consensus"
    repeat_region
    28112. .28361
    /notes="L2 repeat: matches 2243. .2487 of consensus"
    repeat_region
    28389. .28691
    /notes="AluSc repeat: matches 1. .303 of consensus"
    repeat_region
    28694. .28717
    /notes="HSMAR1 repeat: matches 1264. .1287 of consensus"
```

```
repeat_region
    29037. .29333
    /note="AluJb repeat: matches 1. .299 of consensus"
    repeat_region
    29888. .30195
    /note="AluY repeat: matches 1. .308 of consensus"
    repeat_region
    30289. .30601
    /note="AluSc repeat: matches 1. .311 of consensus"
    repeat_region
    30778. .30937
    /note="MER52A repeat: matches 1597. .1755 of consensus"
    repeat_region
    30940. .31306
    /note="MER52A repeat: matches 1096. .1470 of consensus"
    repeat_region
    31307. .31608
    /note="AluSq repeat: matches 1. .303 of consensus"
    repeat_region
    31609. .31902
    /note="MER52A repeat: matches 2. .94 of consensus"
    repeat_region
    31974. .32054
    /note="L2 repeat: matches 2606. .2691 of consensus"
    repeat_region
    32199. .32315
    /note="MER5B repeat: matches 60. .177 of consensus"
    repeat_region
    32401. .32488
    /note="L2 repeat: matches 2452. .2550 of consensus"
    repeat_region
    32903. .33206
    /note="L1M2 repeat: matches 517. .221 of consensus"
    repeat_region
    33241. .34061
    /note="L1M3e repeat: matches 455. .368 of consensus"
    repeat_region
    34739. .35429
    /note="L1M2 repeat: matches 725. .1400 of consensus"
    repeat_region
    35430. .35735
    /note="AluVa5 repeat: matches 1. .306 of consensus"
    repeat_region
    35736. .35873
    /note="L1M2 repeat: matches 1400. .1539 of consensus"
    repeat_region
    35894. .36856
    /note="L1M4 repeat: matches 3844. .4859 of consensus"
    repeat_region
    36857. .37311
    /note="L1M2 repeat: matches 5333. .5787 of consensus"
    repeat_region
    37311. .38137
    /note="L1P13 repeat: matches 3313. .4197 of consensus"
    repeat_region
    38139. .38833
    /note="LTR12 repeat: matches 1. .664 of consensus"
    repeat_region
    38841. .39840
    /note="L1P13 repeat: matches 4195. .5182 of consensus"
    repeat_region
    39843. .39876
    /note="17 copies 2 mer ta 82% conserved"
    repeat_region
    39913. .40870
    /note="L1P13 repeat: matches 5190. .6156 of consensus"
    repeat_region
    40873. .41084
    /note="L1M4 repeat: matches 5769. .5974 of consensus"
    repeat_region
    41085. .41373
    /note="AluSc repeat: matches 1. .297 of consensus"
    repeat_region
    41374. .41623
    /note="L1M4 repeat: matches 5974. .6300 of consensus"
    repeat_region
    41634. .41856
    /note="L1M4c repeat: matches 1811. .1627 of consensus"
    repeat_region
    41943. .42051
    /note="AluSq repeat: matches 1. .109 of consensus"
    repeat_region
    42068. .42344
    /note="L1P4 repeat: matches 5861. .6142 of consensus"
    repeat_region
    42345. .42553
    /note="AluSq repeat: matches 101. .310 of consensus"
    repeat_region
    42688. .43082
    /note="L1M1 repeat: matches 1940. .2717 of consensus"
    repeat_region
    43060. .43547
    /note="L1M9 repeat: matches 2434. .2924 of consensus"
    repeat_region
    43555. .43753
    /note="AluJb repeat: matches 86. .285 of consensus"
    repeat_region
    43754. .43987
    /note="117 copies 2 mer ta 76% conserved"
    repeat_region
    43767. .44004
    /note="7 copies 34 mer 73% conserved"
    repeat_region
    44065. .47403
    /note="L1M9 repeat: matches 2919. .6260 of consensus"
    repeat_region
    48015. .48129
    /note="FLAM_C repeat: matches 30. .142 of consensus"
    repeat_region
    48134. .48314
```

```
repeat_region /rpt_family="Alu"
7297..7325
repeat_region /rpt_family="(CAA)n"
7485..8131
/rpt_family="L1"
10503..10824
/rpt_family="Alu"
10621..10648
/rpt_family="(A)n"
10795..10824
/rpt_family="(A)n"
10877..10970
/rpt_family="(TG)n"
11032..11182
/rpt_family="(TG)n"
11275..12033
/rpt_family="ERV1"
12196..12323
/rpt_family="MER1_type"
12907..13340
/note="similar to EST A1190728 (NID:93741937) qd54h06.x1"
repeat_region 13399..13534
/rpt_family="MIR"
13952..14017
/rpt_family="L2"
14181..14203
/rpt_family="(TGAA)n"
14395..14420
/rpt_family="(GTTTG)n"
14461..14538
/rpt_family="Mariner"
15030..15052
/rpt_family="AT-rich"
15595..15891
/rpt_family="Alu"
15754..15777
/rpt_family="(T)n"
15962..16507
/rpt_family="MaLR"
16835..16953
/rpt_family="MIR"
17029..17180
/rpt_family="MER103"
17670..17790
/rpt_family="ERV1"
18303..18577
/rpt_family="Alu"
19009..19102
/rpt_family="Alu"
19103..19470
/rpt_family="MaLR"
19471..19680
/rpt_family="Alu"
19681..19794
/rpt_family="MIR"
20190..20573
/rpt_family="MaLR"
20880..21004
/rpt_family="MIR"
21787..22090
/rpt_family="Alu"
22975..22999
/rpt_family="AT-rich"
23112..23153
/rpt_family="CT-rich"
23196..23498
/rpt_family="Alu"
24077..24377
/rpt_family="Alu"
25224..25259
/rpt_family="(T)n"
25230..25539
/rpt_family="Alu"
```

```
repeat_region 25897..25946
/rpt_family="(TGG)n"
repeat_region 25973..26158
/rpt_family="MIR"
27003..27077
/rpt_family="L2"
27086..27294
/rpt_family="L1"
27550..27857
/rpt_family="Alu"

Query Match 9.0%; Score 47.8; DB 9; Length 162554;
Best Local Similarity 55.0%; Pred. No. 2.2;
Matches 94; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 26 ATTTATATGCTGATTTATGGGTGATTTTGGTCTCTTATACCTTTATTTATTTCCAA 85
Db 82152 AGTTATGAGTTGACTGTGGGGGTTTGTCTCTTTATGATTTTGTGTTTCCAA 82211

QY 86 ATTTTCTTAAGCAATATTTCTTGTCTAAATCAATAAATTAATCAAAAGAAAAAATCG 145
Db 82212 TTTTCTGCAATATATGTAATCTTCCATATCAAAAAATTAATTAAGTAATAAACA 82271

QY 146 AAAGCAACGCTTGAAGAAAGAAAGTTAGCCCTATCGGGTATATTTGGAA 196
Db 82272 GATGTAGACTAGTCCATATGAAGTCTTGGTCTGGAGACTCTTCATTCTAA 82322

RESULT 31
AL359846 182012 bp DNA linear PRI 16-DEC-2000
LOCUS Human DNA sequence from clone RP11-413C10 on chromosome
DEFINITION 9q31.1-31.3, complete sequence.
ACCESSION AL359846
VERSION AL359846.11 GI:11611367
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Babbage,A.
Direct Submission
Submitted (16-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 9, 2000 this sequence version replaced gi:11557924.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession
numbers given in the annotation table with their source databases:
Emi, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-413C10 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
```

This sequence is the entire insert of clone RP11-413C10. The true left end of clone RP11-31J20 is at 123231 in this sequence. The


```

FEATURES             source
1. .6457
/organism="Dictyostellium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="209116-215573"

BASE COUNT      2444 a      616 c      632 g      2765 t
ORIGIN

Query Match      9.0%; Score 47.6; DB 3; Length 6457;
Best Local Similarity 64.5%; Pred. No. 2.8;
Matches 71; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 61 TCATTATACATTATTATTTATCCCAATTTTCTTAAGCAAAATATTTCTTTGCTAATCAAT 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5600 TCTTTTAGTTTGGGATATTTTAAATTTTGTATAAACAAATAATTTGATTTAAAAA 5541
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 121 AAATTTATCAAAACAAAAAACTGAAAGCAACGCTTGA AAAAAGGAAAG 170
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5540 AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA 5491
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 33
AC123513/C

```

```

DEFINITION
AX4, *** SEQUENCING IN PROGRESS ***.

ACCESSION
AC123513.1 GI:21240650

KEYWORDS
HTG; HTGS PHASE2.

SOURCE
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE
1 (bases 1 to 61052)
Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P.,
Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
Tunggal,B., Cox,E., Quail,M.A., Platzner,M., Rosenthal,A. and
Noegel,A.A.

TITLE
Sequence and Analysis of Chromosome 2 of Dictyostelium

JOURNAL
Unpublished

REMARK
The Dictyostelium Genome Sequencing Consortium

REFERENCE
2 (bases 1 to 61052)
Baumgart,C.

AUTHORS
Direct Submission
Submitted (29-MAY-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

COMMENT
CDS predictions from GeneID may contain errors. Further Information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.shtml)
Funding
Agency : Deutsche Forschungsgemeinschaft (DFG) .
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
Location/Qualifiers
1.. 61052
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="2779865-2840915"
join(448..628,705..803,875..1134)
note="ORF_ID:dd_02708"
/pseudo
/codon_start=1
complement(join(1193..1583,1653..6117,6198..6330))

source

```


SKISALEKHEKTSKVPFLKSKGKIGKNDIEKLLKEDNKLKSKQNOYQIONIKYDKVQ
IQKQLTDDQATAYQQNNKELKLEKSKQDELTIERRKVIKLEQDELTRTPQV
KIPSPGSDTTHLPGANNNNNNNNNNNNNNNGNNMGSSSLSSTNGINGINQV
NNGINGELTHEELIARQSDSRLEARKLEAKKAILKELQIQIDIRVIPERRIQ
NSMPQVQORVOLVSELDIHRNCTKLQNDLAOATIGRRLERALEAFEARHQNI
ERVSQVEVESHELKGEKELVCLIQRNPNVPSQYIOERLLKQKDEKQURKE
IQOLKUDIERYKAPKEEIRYKLIVORDVESKNIENKLEKINTIKQNDKLNNEQ
QLOKSELTLSNLKSKNNININLNDKNIENSKIQQOQQOQQOQQOQQOQQOQQO
KEEELKKEOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQO
KDKEMIDIEKSTENIINNNTNNTEQKLEINIKNIENNNNNNNNNNNNNNNNN
NNNNNDKDIINLEKNEKLEKLELSTVKQLEDITKSEEDYNEIKENKFAOI
KELDITISQNKIQOESQKLEALWEIDSMGKAYQOMLEQNTRLTKQLSDKEDTHAH
LMAENISQOTIRNSKEIQALIEKLNNEKLSQGLMQLKEKSNILQKLSKVT
EDLHSCFPLEKKEKREVRNNAHSLEKTLQDLHLSNLNAELKKKADSLIPALERIDK
AKRLDEKQLLKKLEKATSAANNNNNNNNNNNNNNNNNNSSSEELKSLINQRLCTI
CNDROQNVYIAKCFVCEKCYENIDTRKRCPCSNRAPAETDVHQIYY"
complement (Join (17020, -17639, 17916, -18480))
/note="GeneID exon scores (in order of location ranges):
18,66, 24,28 - GSCJ_ID dd_01628"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM09331.2"
/db_xref="GI:28828362"
/translation="MKLIIIVIFLKLISINGFENIDGNFLIETVETISFNGNFPS
NPFKINGDSSICFSGVGGFNVFSSGLGNSSEIFVXSEEDYFEEDYHETCKHQ
IGITDLVHPNQKLNKNTFFVQEEYYSGCFNLYKNCIDIKDKQNDISYEQIL
KIHSLNISHSTIPSPHIFTFVILVFTFTSNLSFLSLLIINISQWGLISDFGR
FAIFNCALFFSTELGLSAYLVESQFVGEYRNPFSNVGYVLIATQGLTIPFYI
FENKIFSLNIDVOQSFLVKFKIIFSIWILSIPSEIISIFIVDNNPFSLSFIIN
DSPNLIYLLIYIYHPCSKNLIKRLNSYDDGDELSINQSQSQKSIDENLENLY
NK"
complement (Join (19013, -21453, 21643, -22624))
/note="GeneID exon scores (in order of location ranges):
249,99, 106,06 - GSCJ_ID dd_03289"
/codon_start=1
/product="similar to Homo sapiens (Human). Dentin
sialophosphoprotein precursor [Contains: Dentin
phosphoprotein (Dentin phosphophoryn) (DPP); Dentin
sialoprotein (DSP)]"
/protein_id="AAM09332.2"
/db_xref="GI:28828363"
/translation="MTHKKKSLGLFPHSSSSPNLKFLETHEVIEHQOQQOQQNN
NNNNNNHQRPSTTSYSDASSIETSGYLSKTSSSSLSPHLHQNNQNG
NTQNSSGNSNNQOQQOQSIQSSPTSPGKPKKPRLAQPLVNNQNGK
LKKEIELNTKGLNQYTHNLPFNVEDIDNGSGGTSTGNIISHSKSPSSS
SSSSKHKORSPRIPIQSTEDYTDIPRIKMSVEYLFELCLVPGIHSKSNAM
LQRLSOLFEGGDIDISEYDHCIGLLKLPREKPIPIEPYDLHKLRYSVLNEEDS
NIKVSLEISGLKQGLIIRYLIFELLNAVSNVNCVNFMYQNLAICFAPSLIQSDI
SCYDVIENLIENTHSIFGVVVPVVDNNSNVSSSLAHPITIGKSTNVSLLNLTITIL
DHGIDNNSSNNSSSSSSSPYKSPKSPKLNRRNSISPLSSSTSRKSI
NLSNIMISDEYQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQO
RPRKGTVOYLNRTCRPSSWTNNRIKQOHHHHHQOQQOHHQOQQOQQOQQOQQO
NLSLTSPORLNSVGLSEYERGVNRYEYSSLEKSLPKKSLNMQMTVNSNI
INGDEKNSCTSDDEELKISKPIRSTSVNEILKETDNDNNNDQINNNSNNI
PKTITITNTTNTTNNISPIVKSQIITSKAVTPTPTAPMTSSFLSTKQTNPS
SSSSPSTSTSTSSPSSSLSSIDNTMSVNNYRFPANRTPVSPNRFVSPT
TTTTTSGNSFVPRIGKPTPSIPRQSIQCKPNTTSISLSSSSNISKSTSTP
TPTAPLSTLSSVSTSSSKLOEALPTKTTFTQTSPLLPPLSTSTSTST
TST
NNNNNTST
RQLNHTNSSLKSNQIDTPSGINQGMKVQVQKQDPVKKSLAIEKE"
complement (Join (24605, -24844, 24947, -25293, 25414, -25489))
/note="GeneID exon scores (in order of location ranges):
22,64, 31,07, 2,50 - GSCJ_ID dd_01626"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO51012.1"
/db_xref="GI:28828364"
/translation="MVKILPLVLIVLSAVCTTIIGFTYQVYQVYEGVSGSTVTY
YKNKVSFLSFGNDEDTSDYDDVDFEQAOLKFYKAALSMDILAFNLPFSACVM
GMAAFKEISPANOTKLSINGATVFLIATFVLGLPKASREDCEDYAKDSLYTG
SQVDECGGDIYKFTGSHTDGSKLSWGPTTAWIAIVVGLIATFNNIFVYCGPVF
RS"

CDS

CDS

CDS

26170, .27327
/note="GeneID exon scores (in order of location ranges):
24,05 - GSCJ_ID dd_01625"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO51013.1"
/db_xref="GI:28828365"
/translation="MESTSKNGSNLIRKIFINNVGEYILTSDFSPPKVSIEDFIVF
NDLCPDQYQOQFNGNDNNQFYNYIKGREBFKDLGICIVLYLHMFQWENORLNLQ
Query Match 9,0%; Score 47,6; DB 3; Length 136240;
Best Local Similarity 57,3%; Pred. No. 2,5; Indels 0; Gaps 0;
Matches 86; Conservativity 0; Mismatches 64; Indels 0; Gaps 0;
QY 24 TGATTATATGCTGATTATGCGGTGAITTTGCTTCTCTTTATATCTTTATTTATTCCTCC 83
DB 81644 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 81703
QY 84 AAATTTTCTTAAGCAATATTTCTTGTCTATCAATAATTTATCAAAAGAAAAAATCA 143
DB 81704 TTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 81763
QY 144 TGAAGCAACGCTTCAAAAAGGAAGTTA 173
DB 81764 AAAAAATAATTAATAATATATATATAATA 81793
AC006280 163443 bp DNA linear HTG 12-AUG-2000
Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***
AC006280 GI:9797728
HTG; HTGS, PHASE2
Plasmodium falciparum (malaria parasite P. falciparum)
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
Kurdi, O.B., Conway, A.B. and Davis, R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
2 (bases 1 to 163443)
Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurdi, O.B. and Davis, R.W.
Direct Submission
Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
On Aug 12, 2000 this sequence version replaced gi:8810453.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 163443: contig of 163443 bp in length.
FEATURES
source
1. 163443
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clone="3D7"
/clone="PFYAC724"
BASE COUNT 65318 a 15421 c 17069 g 65635 t
ORIGIN
Query Match 9,0%; Score 47,6; DB 2; Length 163443;
Best Local Similarity 50,9%; Pred. No. 2,5;

	GNSLSSTTCTGTSQPLSNITTTKPKQGCKTGKNKGSKSSSPTPYTKVTKRANST EQRTPLLLOKSUTSLKLNNIAVTSDSGIKPFIYSEEFDEDEDGESDESDE DEEEDSEERYBEDDLEDDEDQNYSVDLTKIEXHFLNDEKKNNNNNSGS DETLCOP IAKHQOQQOOQQOVNSXKLNLQSLLDNIPANNNNNQGVKVLSPQ TSQOSTNSPRLFNI FAPENPSFTSPILLHOAT INSS PKVSPPQLSPNGSLSPAG NVSSKRLFFENOITI SSNGTSPKLTNNNNNSNNINNNNSTTTTTTNNNNPJATNIQS ISNNNEKDZDRDKNNNNNDNNNNNNNNNNNNNTSNNNNNNNNNNNNNNNN
	9.0%; Score 47.6; DB 2; Length 61052; Best Local Similarity 57.3%; Pred. No. 2.6;
Matches	86; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Qy	24 TGAATTATATCGTAATTATGGGTGATTTTCCTTCTTTACPTTTATTATCC 83
Dd	6480 TTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTAT 6421
Qy	84 AATATTTCTTAGCAATAATTTCTTGCTAACAATAAAATCATCAAAGAACA 143
Dd	6420 TTATTATTATTAAATATTTTTTTTTTTTATAAATTTTGTATAAAAAAAA 6361
Qy	144 TGAAGCAACGCCTTGAAAAAGGAAAGTTA 173
Dd	6360 AAAATAATATTAAATATATATAATAA 6331
RESULT 34	
LOCUS	AC117070 136240 bp DNA linear INV 12-MAR-2003
DEFINITION	Dictyostelium discoideum chromosome 2 map 2097701-22333941 strain AX4, complete sequence.
ACCESSION	AC117070 AC116976
VERSION	AC117070.2 GI:28828357
KEYWORDS	H7G.
SOURCE	Dictyostelium discoideum
ORGANISM	Dictyostelium discoideum
REFERENCE	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium. 1 (bases 1 to 136240) Glöckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P., Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K., Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and Noegel,A.A.
AUTHORS	Sequence and analysis of chromosome 2 of Dictyostelium discoideum Nature 418 (6893), 79-85 (2002)
TITLE	The Dictyostelium Genome Sequencing Consortium
JOURNAL MEDLINE PUBLISHED	2 (bases 1 to 136240) Baumgart,C.
REMARK	Direct Submission
AUTHORS	Submitted (06-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
TITLE	(bases 1 to 136240) Baumgart,C.
JOURNAL	Direct Submission
REFERENCE	Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
AUTHORS	(bases 1 to 136240) Baumgart,C.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT	On or before Mar 4, 2003 this sequence version replaced gi:20042931, gi:20066187. CDS predictions from GenBank do not necessarily reflect true genes. Further information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I (http://www.uni-koein-de/dictyostelium/project.htm#funding) Agency : Deutsche Forschungsgemeinschaft (DFG).
FEATURES	Location/Qualifiers

[illegible]

1. 136240

/organism="Dictyostelium discoideum"

/mol_type="genomic DNA"

/strain="AX4"

/db_xref="taxon:44689"

/chromosome="2"

/map="2097701-2233941"

3052..5372

/note="GeneID exon scores (in order of location ranges):

223..19 - GSCJ_ID dd_03123"

/pseudo

/codon_start=1

join(5982..6435,6513..8440)

/note="GeneID exon scores (in order of location ranges):

11..34, 121..54 - GSCJ_ID dd_01623"

/codon_start=1

/product="similar to Plasmodium falciparum. Hypothetical protein"

/protein_id="AA009325.2"

/db_xref="GI:28828358"

/translation="MKKTIIFLFTIILIKLSNDRLVGVGRNNKQSFINKNTWYF
EKVPTKDSLLSNGGLLIDYEITCSSYFTINNGSFYLVGVGVVDSMDINRGSFI
MENGALINVNNSIYVYKYFEINNSIENPTEKNTTSTGNFIIGFLNIYOKSLFLN
CNKLEINNGSINVENNSKININGGLIMKDKSQIYLNNSNIEINGDAMLNQSIFNTM
NNLDFISGLHMLDLSUFLFDNMNETINGDLLLISGLFVNDKXSLIYIGDLKIN
GNSGFTIGMESIINIGDFYSVNSFEITKISFNWSGSAEFGIKNFSCDFNSHK
STFNSTDIYVVIDSVFKSHGLSTVNNKSLIAQNSIYVSDLLIINGEPTNTFTYV
IEGNSILQNSYINLSDIPIFINKNLIIPHNSYINLNSKFIPLDNGIVKTSGLNVL
NNSTFEINQNGIYNLSSGIYTYNDTHIKSNNTILKNGDIFINDNVEINSPFIN
GDGNGSGSGSGGGNDSLQTLNNHNTIYIYEFNENGSENLNGTLLSDNSIQ
LNGSIKINGNINGTIYIGDFEGSGFVYNNFIIGNFIQNSDSSSKIIIVIDSLFY
STINISKLADINGIYIKINELNLLNNNNNNNNNNNNNTSSSSSTNEKTILKLVN
FDSKSVNGFNKIVATYSGKNEKELSSPPCVDARVSTSPVLIDVNDCFPNDKV
NLSKSGAIVGVIACAIITAILMSIVHYRERARFLGMLAKAKLSKISKLSRK"

join(8760..8936,9059..9136,9233..9566,9646..10421)

/note="GeneID exon scores (in order of location ranges):

14..86, 5..34, 37..59, 71..71 - GSCJ_ID dd_01622"

/codon_start=1

/product="hypothetical protein"

/protein_id="AA009326.2"

/db_xref="GI:28828359"

/translation="MSHANAIIWYLLDGAWKPYDSSLIDGFEKSPNSSLIATGNKKI
KVDNRYEINTPIADILNSFNNSINDPLVGIQRFPDAMGRARVKTANTTTSFDD
DLOAFYNSQPKQKTNEIKGIKICGKSGVDTMTKTDYLLILKPSDLDDYKTFNDI
ISKAELGIICVESTWIDSCIKSKLVDPHTAFVGOKTAPVVAIPNTTTTATTATT
TNTGTATSPKSKNNSTTNVLVNSGFLIKGSEWAGYCSSDDHHYPLVWTVDD
FVGGSEGVITWINLGAKTYKGVAVNGEINGFQBYEISGEDDVEVPNDYKT
IYGDITIGKVGDSYFKLKTSPVPTNPQDSQWGTSTSIENFKLDVVKRDEGNTG
TITWGSYDDAIATFTGKITSTSEVSDYKGEKGLGPBLSHPFKLIPSQQSDHSIKF
RIEK"

join(11103..11133,11397..11698)

/note="GeneID exon scores (in order of location ranges):

-2..92, 19..55 - GSCJ_ID dd_01630"

/codon_start=1

/product="similar to translation initiation factor 3 (eIF3): Sui1p [Saccharomyces cerevisiae]"

/protein_id="AA051010.1"

/db_xref="GI:28828360"

/translation="MSSIQNLSADPFAEDSNLGGDSGIGSYVHIRIFQRSGKAVTV
EGSPKPTLDKILKHLKSLNCNGNIVBEENGLMIKLQGDKKRKETALFIEBKISOK
PSIKHGK"

join(12993..13277,13378..13450,13554..13748,14017..14609,
14702..16798)

/note="GeneID exon scores (in order of location ranges):

20..39, 2..30, 10..80, 63..82, 215..25 - GSCJ_ID dd_01629"

/codon_start=1

/product="similar to Homo sapiens (Human). Hypothetical protein FLJ12238"

/protein_id="AA051011.1"

/db_xref="GI:28828361"

/translation="MSMDHQDLKRRDPADSSQRPSSPKKISISSSEASVPIGGPPIDQ
NLILIFNRKAMRVVEQKIETNDRSQIKQLNTHIKYKENCISCLCRVQDLNSGLDL
LINRVDFEMANRLLPKONIYENRQKSYITEPVTLYDKYTLDLSLOKKVQTSFIF
L

FEATURES


```

Matches 113; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 26 ATTATATGCTGATTTATGGGTGATTTTGGTCTCTTTTATCTATCTTTTATTTATTTCCCAA 85
Db 37449 AATTTTATGTTGATGATTCATTTAAACACTTCACCTTTTCTTTTCTTTTCTTTTCTTTT 37390
QY 86 ATTTTCTTAAAGCAATATTTCTTCTGCTAATCAATAAATATCAAAAGAAAAAAACTG 145
Db 37389 TTTTCTTTTACAAATATTTAAATTAATATACATAATTTATTAATAAATAAACAACAA 37330
QY 146 AAGCAACGCTTGAAGAAGGAAGTAGCCCTATCGGTATATTTTGAAGTTCTGATAAA 205
Db 37329 AAAACAATAATTTATAACAATAAATAAACAACAGTTATTTTATTTATTTATTTATTT 37270
QY 206 TACTACRTGTTCTCTTCTAAAGTCCCACTCTCTGTTTCTTT 247
Db 37269 TTTTCTTTTGTACTTTTATATGAACTTCAATGTAATAT 37228

RESULT 36
AC004709 196149 bp DNA linear HTG 01-APR-1999
DEFINITION Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
AC004709
ACCESSION AC004709.3 GI:4558585
VERSION HTG: HTGS PHASE1.
KEYWORDS Plasmodium falciparum (malaria parasite P. falciparum)
SOURCE Plasmodium falciparum
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 196149)
AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Tanaki,T., Kurdi,O.B., Conway,A.B. and Davis,R.W.
TITLE Plasmodium falciparum 3D7 chromosome 12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 196149)
AUTHORS Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT On Apr 2, 1999 this sequence version replaced gi:4337173.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 47631: contig of 47631 bp in length
* 47632 47831: gap of unknown length
* 47832 179129: contig of 131298 bp in length
* 179130 179329: gap of unknown length
* 179330 196149: contig of 16820 bp in length.
FEATURES
Location/Qualifiers
source
1..196149
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"

BASE COUNT 80057 a 19753 c 18900 g 77138 t 401 others
ORIGIN

Query Match 9.0%; Score 47.6; DB 2; Length 196149;
Best Local Similarity 50.9%; Pred. NO. 2.5;
Matches 113; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 26 ATTATATGCTGATTTATGGGTGATTTTGGTCTCTTTTATCTATCTTTTATTTATTTCCCAA 85
Db 62724 AATTTTATGTTGATGATTCATTTAAACACTTCACCTTTTCTTTTCTTTTCTTTTCTTTT 62783
QY 86 ATTTTCTTAAAGCAATATTTCTTGTGCTAATCAATAAATATCAAAAGAAAAAAACTG 145

TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
REFERENCE 1 (bases 1 to 250029)
AUTHORS Gardner,M.J.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
source
1..250029
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="10"
repeat_region
81..109 /rpt_type=tandem
repeat_region
476..506 /rpt_type=tandem
repeat_region
688..718 /rpt_type=tandem
repeat_region
690..733 /rpt_type=tandem
repeat_region
697..757 /rpt_type=tandem
repeat_region
945..1024 /rpt_type=tandem
repeat_region
965..1024 /rpt_type=tandem
repeat_region
1240..1274 /rpt_type=tandem
repeat_region
1289..1319 /rpt_type=tandem
repeat_region
1435..1483 /rpt_type=tandem
gene
<1560..>2471
locus_tag="PF10_0059"
mRNA
join(<1560..2109,2341..>2471)
locus_tag="PF10_0059"

```

```
/locus_tag="PFL0935c"
complement(join(11613. .12941,13715. .19156))
/locus_tag="PFL0935c"
/notes="HMPfam hit to PF03011, Plasmodium falciparum
erythrocyte membrane protein (PFEMP), score 1.3e-56"
/codon_start=1
/product="erythrocyte membrane protein 1 (PFEMP1)"
/protein_id="AAN36275.1"
/db_xref="GI:23496720"
/translation="MAAAAGGGSSODAKHVLDEIGKIVYETVEKDAKLYKEALGNL
QKAKNGERASSNKTCTLYKEYEHFNGDANSNRYECKELSGKMGENRSDTLGGCT
NKIEGNKNNGCAYRRLHLCHLNLESIDTTSITDILLAEVCYAAKPEGETLITQ
HOHQOQTNGTASQCLCTVLARSADIGIVRGKDLFNGDEEKKRDELEENLITQ
KIHSGTSTRGQNGLRKYKGDANKFYQLREDDWDRNHTVWEAITCGAGESDKY
FROTACTGTTHKQRCDDKANVDPPTFYVQYLRFWEAEWDFCRKKNKGLKDV
LNCRCGSDTDRYVCSGNGDCTOTIRAIKLVIGEHCTNCVSRKCEBECKCECFQK
LKNQKTEISGGGASGASGRSRTAATIKYEGYKFEYEEPKKNGYGTVDLFL
GLNNKACKDINEEKIDFTKNVEDDKNINKEGTFYHSQYCPGCGMKRTKNGN
GWEEDTNCNKLKYPTSAKTEIKLKSGEKHDDIKELIEFCQTKQNGDSGGGG
NSEKKELYDMWKYQIGELTKDQEGMDPDYDLVETGGGLCILENKKKSKDPKS
QNEPEQKTFNEFFYFGRFLNDSMYRWKGVNSCINKSKREKCEBECKCECFQK
WIGKKKEWDKIKHEFTKODFGNDALIGPEMKSPSVLKFVLNIDELFONIKSGYGVN
KELKGINLDERKKEEERADGVVGASGDNQNTTIDKLLDHEDKATKCRNQ
EPTKASKDELARSASDTPREDSPADLDDDEHDEDEDDDAESEESEKKEE
EEHKEEENAEPEEKESGSSPEEKEKAPAGPACNIVQTLFSIDKFSDAC
TLKYGTATPGWKCVTPSGDKAATSESGSSDGAGGVNGGALQRNKRDLATPS
AKSGDTGGKDGATGSDGSIQVPPRRRLYIKLHDWAEKVGDTATQPOVDTPSOD
KLURDAFIESAAVETFLWHRYKKIKDEKLEQORORENGELPGLSSGDSNDPQS
KLEKGIPOGFLRQMPYTLADYKDLIYSGSNDNTKSTYNDIINGDKETQERKDIKG
ALSTYNSGDTTPPVSIVTPTQWELHGPHIWNGMICALTYKENGIDKPEVDTA
VRAQLWDGNTSPONKQYQTNVKBEEASGAKINIPDYSGDNTPLTNTPLKNFVEI
PTPFWLHWSGDFQQRKRMKNVKNCRNIRERGHEYCGDGHDTPTLQHNHML
ADPDQDCHIQCKYKRWIEKKEPEFHQNKYKQKGLNGNSNDGNTKPFKEIEN
RSPDQFLSLHCKNNGSGDPKNKIDFKEPLKTRGPLEYCKTCPPNVNCGNSRR
SGNDQCTAVNGEMEKIPENGNSTTIDVHMLDRGPYMEKKSOKLSENPLKT
SRLEGIREQNTCKKADKNIDVCKLQDFQVLDNQYITPFRVLLIWLDEFIEGYL
LKRRKIIETQKEGTCDDESKDCACQKQVKEKKTWEQIKHEFRNRQKQSGDN
DMKSSVRQLDPLIYRMDLANGKINELKFLSYECKCVNAGSEKDVVCECLQK
LETKAACKDPSGSENGKTPCQBYTPDDDDLEENETVKAFTCPKVPKKEQTE
ETCDKADEKVEBTAQNEENIPAGAPPADSPAGPATDSGKENVPEPPAPAP
PSPPIPLASTLSKVGIGFAAFTYFLKKKTASVGNLFOILOIPKSDYDIPTLKS
KNYIYFRSGTYKGTIIMEGDADSKYTFMSDITDITSSSEBELDNLITVPGS
PKYKTLIEVLEPSGNTTASGKNTPSDTONDIONDGPSSKITDNEWNLKODEFLN
MLQNKQKVPNDYSSGIDIPNTQNTLYFDNNOEKFITSIHDRNLTYGEESYNVNM
VNDNIPINRDNNTYPSGIDILINDSLNSKVDIYDELLKREKNELFGEHHPKHNTYN
VAKPARDPTTNOINLFHKLDRHDMCEKMDTNNKVDIILNQLKESWENDNSNGKNT
SGNITPSIDPSKLSDLSTNKLMSDSVSIQIHIDKPNQVDDNIYLDYTPPKYTVDN
INPVDTHNPLNPNVDQNSNLTFPSNPNPAYONIYIDHNEDLPKVKQIEMSVK
NGEMAKEK"
repeat_region
12602. .12648
/rpt_type=tandem
repeat_region
12960. .13070
/rpt_type=tandem
repeat_region
12966. .13039
/rpt_type=tandem
repeat_region
13048. .13072
/rpt_type=tandem
repeat_region
13469. .13493
/rpt_type=tandem
repeat_region
13562. .13586
/rpt_type=tandem
repeat_region
16563. .16614
/rpt_type=tandem
repeat_region
17915. .17943
/rpt_type=tandem
repeat_region
19242. .19283
/rpt_type=tandem
repeat_region
19334. .19422
/rpt_type=tandem
repeat_region
20141. .20273
/rpt_type=tandem
repeat_region
20940. .21079
/rpt_type=tandem
```

```
Query Match          9.0%; Score 47.6; DB 3; Length 252650;
Best Local Similarity 50.9%; Pred. No. 2.4;
Matches 113; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 26 ATTATATGCTGATTTATGGGTGATTTTGCTTCCTCTTTATACCTTTTATTATTCCTCAA 85
DB 90609 AATTTTATGTTGATGTAATTCATTTAAACACCTTCACCTTTTTTTTTTTTTTTTT 90668

QY 86 ATTTTCTTAAGCAATAATTTCTTTGCTAATCAATAAATATCAAAAGAAAAAAACTG 145
DB 90669 TTTTTTTTTTCAAAATATTAATAATTAATATACATAATTAATAAATAAACAACAA 90728

QY 146 AAAGCAACGCTTTGAAAAAGGAAGCTTAGCCCTATCGGCTATATTTTGGAAAGTTGTA 205
DB 90729 AAACAATAAATTTATAACAATAAATAAACACCACTTTATTTTATTTATTTATTTAT 90788

QY 206 TACTACRGTGTTCTCTTAAGTCCCACTCTCTGTTCTTTT 247
DB 90789 TTTTCTTTTGTACTTTTTTATATGAATCTTCAATGTAAATAT 90830

RESULT 39
PFMAL1P2_1
WPCOMMENT
Sequence split into 4 fragments LOCUS PFMAL1P2 Accession AL031745
Fragment Name      Begin      End
PFMAL1P2_0         1      110000
PFMAL1P2_1         100001   210000
PFMAL1P2_2         200001   310000
PFMAL1P2_3         300001   384550
Continuation (2 of 4) of PFMAL1P2 from base 100001 (AL031745 Plasmodium falciparum DNA f.1)

Query Match          9.0%; Score 47.4; DB 3; Length 110000;
Best Local Similarity 52.8%; Pred. No. 2.8;
Matches 102; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 28 TTATATGCTGATTTATGGGTGATTTTGCTTCCTCTTTATACCTTTTATTATTTCCCAAT 87
DB 36043 TAATATATTATATAATATATATTTTTTTTTTTTTTTTACAAAATTTCCAGTAAGCAAAATA 36102

QY 88 TTTTCTTAAGCAATAATTTCTTTGCTAATCAATAATATCAAAAGAAAAAAACTGAA 147
DB 36103 TATCCATTCGTATACAAATCTTTTCTTAATATACACATATGAACAAAAAAAAGAA 36162

QY 148 AGCAACGCTTGAAAAAGGAAGTTAGCCCTATCGGCTATATTTTGGAAAGTTGTAATAA 207
DB 36163 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 36222

QY 208 CTACRTGTTCTCT 220
DB 36223 TTATCTATACTTTT 36235

RESULT 40
AX347077/c
LOCUS
DEFINITION
Sequence 2148 from Patent WO0200928.
ACCESSION
AX347077
VERSION
AX347077.1 GI:18494965
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
artificial sequences.
REFERENCE
1
AUTHORS
Olek,A.; Piepenbrock,C. and Berlin,K.
TITLE
Diagnosis of diseases associated with the immune system
JOURNAL
Patent: WO 0200928-A 2148 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
source
1. .113515
/organism="synthetic construct"
/mol_type="genomic DNA"
```

```
Qy 24 TGATTTATATGCTGATTTATGGGTGATTTTGCTCTCTTTATATCTTTATTTATTTATCC 83
Db 101048 TAATTTTATCATATTTATTTATTTTATATATATTTATTTTGGTTCATTTTGTGA 101107
Qy 84 AAATTTTCTTAAGCAATATTTCTTGTGTAATCAATAAATATCAAAAGAAAAAACAAC 143
Db 101108 ATATTCATCTTATTTTATTTATATATTTTTCATGATAGATTTTAAAAAGAAATAAAAA 101167
Qy 144 TGAAGCAACGCTTGAAAAAGCAAGTTAGCCCTATCGGCTATATTTTGGGAAGTTGTAA 203
Db 101168 ATAATACTTTGTTAAATAATGTAACGATTTATAAATATGTAATATGCAATATATAAA 101227
Qy 204 AATACTACTGTCTCTCTCT 223
Db 101228 AATTTAATATACATCATCT 101247

RESULT 38
LOCUS AE014847
DEFINITION Plasmodium falciparum 3D7 chromosome 12, linear INV 11-FEB-2003
complete sequence.
ACCESSION AE014847
VERSION AE014847.1
KEYWORDS GI:23496718
SOURCE plasmodium falciparum 3D7
ORGANISM plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 252650)
Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
Anguoli,S., Perlea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
MEDLINE 2255705
PUBMED 12368864
REFERENCE 2 (bases 1 to 252650)
Hymen,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
Direct Submission
Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
3 (bases 1 to 252650)
Hymen,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
Direct Submission
Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES
source
1..252650
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="12"
repeat_region 90..134
/rpt_type=tandem
repeat_region 96..134
/rpt_type=tandem
repeat_region 586..669
/rpt_type=tandem
repeat_region 724..748
/rpt_type=tandem
repeat_region 1220..1250
/rpt_type=tandem
repeat_region 1331..1363
```

```
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
gene
mrna
CDS
```

```
/rpt_type=tandem
1440..1476
/rpt_type=tandem
1758..1801
/rpt_type=tandem
1907..1971
/rpt_type=tandem
2150..2184
/rpt_type=tandem
2301..2344
/rpt_type=tandem
2484..2513
/rpt_type=tandem
<2608..>8601
/locus_tag="PFL0930w"
<2608..>8601
/locus_tag="PFL0930w"
2608..8601
/locus_tag="PFL0930w"
/note="HMPfam hit to PF00637, 7-fold repeat in Clathrin
and VPS, score 2.1e-24"
/codon_start=1
/product="clathrin heavy chain, putative"
/protein_id="AAN36274.1"
/db_xref="GI:23496719"
/translation="MSQNNPLSCVADNLINVDIQNESFRIGNVSVEGDKYICVKNV
NENTQVVVNLNKNSTKMKMAESVLIHPNDPILALKGTIKNMTIPIQVFNLETKE
KICSLNENYNNYKWINNDTIALVCEKNVYHNIDHTKKNKNDNNNNNNNNNN
NNNNNNNTLTQVFERKAQPIDNNSQLIYSTDKMKWCLCGISTODQSGSIDGYM
OLYCECKLHTQTEGTCFSGFVDNLMKPLFCFVEKKNSISRLHMLDIYNTK
ETTPYIKVEINENINENLDPYIYISLNTLQGVYIYTKGSYVVFDEGTLISKE
KISDNIPICDNSKINGIVAVNKKGIIFITINVFLINHLNLSNEFKDKLIKLC
VKYGPCCDXTISAVKCCINDMDFKASKILCLMKNTKIPEHNSNIAESILMKNKI
DIRIIPRLTQQULNSFKSPKTSQGLSPLLLYSVLLEVKLTLYTESIELKPVVLQ
KKBYLEKWIKDDKLTCEBGLDVLKVDURLALNIYLRCSHNKILISTYCLLNMFN
VLSYINFPKQINFQYIGVIFITINVEFPQGNADNKSFAFGVNSGSKNETMDFNDG
AASISNQDMS INKLNGHINEVAIQIKFLCENNI SF DINKI IDYLLSKKLOBEAT
ILLDLKENKAHEKNLQTLPEFLNYYNNOVAETLFQMDI FTYYDNKNIAYLCBEKGI
YORALENTYINDIKRVITKSTCFQKGNNTNTTTTTSNMGDGFDMNVSKGI
SIEWIKYFSTLSDSVCEELLDFPMKSKINMEVVISICVQYDKIGIKKILNPFEN
KNYEGIFPYFYNILNLPNLIVKNNNDNNSLSSTYNNNDSDASILLTSDICSTP
QYSYETANLKLIEDVHYIMFYIEACVKNINIQLDRI CKDKNAKNPQIKNFKLCK
LSDPRLIYVCDIHNYIEELAEYLYKNSLLKYIEVYVIKVPNNNAHVIGVLLDDAS
EDFLNLNNIKNISNIGLIEIAEKRLKLLLPWLESANEGVENIELNALAKYI
IDLKNPENFLKNNFYDKMWIKYCEDLDPHLAYAYERSGSCDEELIYITKNGLI
FKLQARVLSRQSLKHWLVLDENKRYKRVIDQVIGSTLIESNABEITVTVAFIE
KKUSSELIELLEKIVLHNSFSDNNKLNQLLILTAIKSDSKKVMYENIRLNDVSGPQI
AAVAYEKLREAEFVYIKFCYTSAISVLLDKILYNNKQKNQNFHFSHSDSTVPYI
SNNNKDNIPGNKSDNENDEYLKQFTQNNSYREKYGYNNNDISNNSNNIHLHPEF
NNEKFSDDTLHNYEDIDSTESTCDLSNYEDLNRAIEFAKCNNDVWFILGKAOLK
LNKIDSIDSPVKSNPNPEAYKEVLECKENNFYELLITLNTLREQNSLKDVLVDSGL
LYAYAKUKKNNMKFINTNSANLQIGRLYKEBEYEVAKILYSNIPNNKLTACYI
LKKKEYALNLEAKKAKSLTKWEVNFICVKYKQKYAHTAGLOLMHADHLEIKI
YERKKYIENLNLNENLNNERAHVIGYTELGIYAKYKPEKLMFETRYNNKNMTRK
LIDVCEYELLKEAVLYISYDENLAVDTIIKHSPTAYTADTFQVHKVNTSIIH
KVIDFYIEEHPNLNLLKILENKIDNNRLVQTLKSNLPLIOKYLEIDAOQNTIAH
NETLINEYLNDDYISLRNSIDEDNFQNTNLNKLHNKLPERRIAALYKKNKY
KEATNLSKKEKQKDAIEARVSKNNVTIEDLINFYIESKKNKEAFACCLIVCYDLLKP
DYILEIVTSGFGKQDAMLFYIISDYNTQIEHMKKQIIEDEKESKKNKNSPNDYSAM
NNQFVSLNNLSIMPPQNNFMSSNSFDKYDMFNNTHF"
3056..3105
/rpt_type=tandem
4923..4948
/rpt_type=tandem
8613..8642
/rpt_type=tandem
8716..8742
/rpt_type=tandem
9080..9130
/rpt_type=tandem
complement(<11613..>19156)
/locus_tag="PFL0935c"
complement(join(<11613..12941,13715..>19156))
```


chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTA-250D13 is from a release of the human BAC library CITB-HS-A. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
VECTOR: pBeloBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The actual start of this clone is at base position 1 of RG250D13, actual end is at 176109 of RG250D13.

This clone contains STS sWSS2794 (NID:gl113588) and sWSS350 (NID:g484309).

FEATURES

Source

Location/Qualifiers

1..176109
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q31"
/clone="CTA-250D13"
/clone_lib="CITB-HS-A"

misc_feature
5770..5950
/notes="similar to EST T97069 (NID:g735693) ye50f03.r1"
misc_feature
5770..5949
/notes="similar to EST AW490098 (NID:g7060333)"
misc_feature
6407..6665
/notes="similar to EST W21257 (NID:gl298308) zb58b07.r1"
repeat_region
14504..14805
/rpt_family="L2"
repeat_region
14806..15107
/rpt_family="Alu"
repeat_region
15108..15194
/rpt_family="L2"
repeat_region
15360..15462
/rpt_family="MIR"
repeat_region
15633..15723
/rpt_family="L2"
repeat_region
15790..16074
/rpt_family="L2"
repeat_region
16787..16890
/rpt_family="MIR"
repeat_region
17143..17416
/rpt_family="MER1_type"
repeat_region
18425..18528
/rpt_family="MIR"
repeat_region
18853..18989
/rpt_family="Alu"
repeat_region
19024..19143
/rpt_family="Alu"
repeat_region
19338..19482
/rpt_family="MIR"
misc_feature
19664..20147
/notes="similar to EST AA034047 (NID:gi1505856) zi05e04.s1"
repeat_region
20436..20733

repeat_region
20898..21735
/rpt_family="L1"
repeat_region
21736..22031
/rpt_family="Alu"
repeat_region
22032..22079
/rpt_family="L1"
repeat_region
22690..23308
/rpt_family="L1"
repeat_region
25605..25729
/rpt_family="L2"
repeat_region
26006..26319
/rpt_family="Alu"
repeat_region
26377..26788
/rpt_family="L2"
repeat_region
28294..28390
/rpt_family="MIR"
repeat_region
30562..30741
/rpt_family="MIR"
repeat_region
30945..30984
/rpt_family="L2"
repeat_region
31078..31146
/rpt_family="L2"
repeat_region
32776..32830
/rpt_family="L2"
repeat_region
33004..34302
/rpt_family="L1"
repeat_region
34389..34576
/rpt_family="L2"
repeat_region
35613..35812
/rpt_family="MIR"
repeat_region
35980..36553
/rpt_family="L1"
repeat_region
36887..37180
/rpt_family="Alu"
repeat_region
37807..37949
/rpt_family="Alu"
repeat_region
38095..38581
/rpt_family="MaLR"
repeat_region
38876..39185
/rpt_family="Alu"
repeat_region
39402..39485
/rpt_family="L2"
repeat_region
39826..39926
/rpt_family="MIR"
repeat_region
41270..41572
/rpt_family="Alu"
repeat_region
41637..41939
/rpt_family="Alu"
repeat_region
42310..42463
/rpt_family="Alu"
repeat_region
42583..42791
/rpt_family="MER1_type"
repeat_region
43575..43952
/rpt_family="CRI"
misc_feature
44284..44598
/note="similar to EST AW946860 (NID:g8124633)"
repeat_region
44407..44534
/rpt_family="Alu"
repeat_region
44582..44701
/rpt_family="CRI"
repeat_region
44716..44926
/rpt_family="CRI"
repeat_region
47256..47583
/rpt_family="L2"
repeat_region
47635..48161
/rpt_family="L2"
repeat_region
49690..49883
/rpt_family="L2"
repeat_region
50136..50338
/rpt_family="L2"

Query Match

9.0%; Score 47.4; DB 9; Length 176109;

```

/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT      35172 a      1174 c      22520 g      54649 t
ORIGIN

Query Match
Best Local Similarity 9.0%; Score 47.4; DB 6; Length 113515;
Matches 93; Conservative 1; Mismatches 77; Indels 0; Gaps 0;

Qy 55 CTTCTCTTTTATCTTTTATTTCCCAATTTTCTTAAGCAAAATTTCTTTGCTA 114
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41662 CTTCTCTTTTATCTTTTATTTTCTTTTAAACAACTTTTATCTTCTT 41603
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 115 ATCAATAAATTATCAAAAGAAAAAACTGAAGCAAGCTTTGAAAAAGGAAAGTTAG 174
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41602 ATTATATATTATCACTATAAACTCATAAATCTAAATATTAAATAAATCA 41543
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 175 CCTATCGGTATATTGGAGTTGTAAATACTACRTGTTCTCTCTAA 225
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41542 TTCCCACTATTATTATTTTATATCCAAATTTATCAAAATTTAACTTAA 41492
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 41
AC004153
LOCUS      156060 bp      DNA      linear      HTG 12-AUG-2000
DEFINITION Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 2 unordered pieces.
AC004153
VERSION    AC004153.7 GI:9797733
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Plasmodium falciparum (malaria parasite P. falciparum).
ORGANISM   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE  1 (bases 1 to 156060)
AUTHORS    Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tanaki,T.,
Kurti,O.B., Conway,A.B. and Davis,R.W.
TITLE      Plasmodium falciparum 3D7 chromosome 12
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 156060)
AUTHORS    Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
TITLE      Direct Submission
JOURNAL    Submitted (18-FEB-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT    * NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 102169: contig of 102169 bp in length
* 102170 102369: gap of unknown length
* 102370 156060: contig of 53691 bp in length.
FEATURES
source
1..156060
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5933"
/chromosome="12"
/clones="PPVAC812"
/clone="3D7"

BASE COUNT      62615 a      14889 c      15137 g      63219 t      200 others
ORIGIN

Query Match
Best Local Similarity 9.0%; Score 47.4; DB 2; Length 156060;
Matches 99; Conservative 1; Mismatches 87; Indels 0; Gaps 0;

Qy 39 TTTATGGTGATTTTGGTCTCTTTATACATTTTATTTATCCAAATTTCTTAAGC 98
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 69975 TATATCGCTACTGAAGTATGCCTATTATTATCTATTATTATGTTTCAAAATATTTTATAT 70034
Qy 99 AAATATTTCTTTCCTAATCAATAAATTATCAAAAGAAAAAACTGAAGCAAGCTTGT 158
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70035 GGATATAATATTTTGGCGCTAACAATAGAAAAAGAAAAAAATTTTCAATTTCT 70094
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 159 AAAAAGGAAAGTTAGCCCTATCGGTATATTTTGGAAAGTTGTAAATACTACTACRTGTTCT 218
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70095 ATTATAAAATATATTATTATTGTTTAAATGTTATTATTTTATAAAATTTTCAATTTCT 70154
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 219 CTTCTAA 225
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70155 ATTATAA 70161
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 42
AC003992/c
LOCUS      176109 bp      DNA      linear      PRI 04-JUN-2002
DEFINITION Homo sapiens BAC clone CTA-250D13 from 7q31, complete sequence.
AC003992
VERSION    AC003992.2 GI:21322221
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 176109)
AUTHORS    Sulston,J.E. and Waterston,R.
TITLE      Toward a complete human genome sequence
JOURNAL    Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE    99063792
PUBMED     9847074
REFERENCE  2 (bases 1 to 176109)
AUTHORS    Minx,P., Hinds,K., Sutterer,C., Becker,M. and Ozersky,P.
TITLE      The sequence of Homo sapiens BAC clone CTA-250D13
JOURNAL    Unpublished (2001)
REFERENCE  3 (bases 1 to 176109)
AUTHORS    Waterston,R.
TITLE      Direct Submission
JOURNAL    Submitted (14-JAN-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE  4 (bases 1 to 176109)
AUTHORS    Waterston,R.
TITLE      Direct Submission
JOURNAL    Submitted (03-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE  5 (bases 1 to 176109)
AUTHORS    Waterston,R.
TITLE      Direct Submission
JOURNAL    Submitted (04-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE  6 (bases 1 to 176109)
AUTHORS    Waterston,R.
TITLE      Direct Submission
JOURNAL    Submitted (04-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT    On Jun 4, 2002 this sequence version replaced gi:2772533.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
-----
Center project name: H_RG250D13
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate

```

```

repeat_region 8606..9201
/rpt_family="L2"
repeat_region 9202..9391
/rpt_family="MER1_type"
repeat_region 9392..9678
/rpt_family="L2"
repeat_region 9679..10204
/rpt_family="ERV1"
repeat_region 10205..10574
/rpt_family="L2"
repeat_region 10575..10593
/rpt_family="MIR"
repeat_region 10594..10901
/rpt_family="Alu"
repeat_region 10902..10937
/rpt_family="MIR"
repeat_region 10938..11169
/rpt_family="L1"
repeat_region 11170..11249
/rpt_family="MIR"
repeat_region 11289..11492
/rpt_family="L1"
repeat_region 11972..14123
/rpt_family="MER1_type"
repeat_region 14124..14359
/rpt_family="MER2_type"
repeat_region 14360..14416
/rpt_family="MER1_type"
repeat_region 14417..14708
/rpt_family="Alu"
repeat_region 14709..14756
/rpt_family="MER1_type"
repeat_region 17593..17812
/rpt_family="Alu"
repeat_region 18269..18834
/rpt_family="MER99"
repeat_region 19429..19573
/rpt_family="MER1_type"
repeat_region 19579..19647
/rpt_family="MER99"
repeat_region 19700..19817
/rpt_family="MIR"
repeat_region 20224..20267
/rpt_family="MER1_type"
repeat_region 20268..20570
/rpt_family="Alu"
repeat_region 20571..20700
/rpt_family="MER1_type"
repeat_region 29065..29261
/rpt_family="L2"
repeat_region 29262..29555
/rpt_family="Alu"
repeat_region 29556..29941
/rpt_family="L2"
repeat_region 29942..30131
/rpt_family="MER1_type"
repeat_region 30132..30316
/rpt_family="L2"
repeat_region 30335..30633
/rpt_family="Alu"
repeat_region 31082..31204
/rpt_family="MaLR"
repeat_region 31219..31329
/rpt_family="MaLR"
repeat_region 31422..31509
/rpt_family="MaLR"
repeat_region 32196..32302
/rpt_family="MaLR"

```

Query Match 9.08; Score 47.4; DB 9; Length 176379;
 Best Local Similarity 53.68; Pred. No. 2.7;
 Matches 96; Conservative 1; Mismatches 82; Indels 0; Gaps 0;

```

QY 34 GCTGATTTATGGGTGATTTTCTTCTTTATATCTTTTATTTATTTTATTTTCTCCAAATTTTCT 93
Db 22714 GATAATTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 22655
QY 94 TAAGCAATATTTCTTTGCTATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 153
Db 22654 AAAAAACAAATTTTCCACAGAAATCATACATTTTTCAAAAATATAAACAATATAAAGTGTAA 22595
QY 154 GCTTCAAAAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAATAATCTACTAC 212
Db 22594 GATTTTCATATAACAGATATTTCCAAGAGACTATTTTGAAGTGGGAATATTTTACA 22536

RESULT 44
AE014848
LOCUS 250707 bp DNA linear INV 11-PEB-2003
DEFINITION Plasmodium falciparum 3D7 chromosome 12, section 5 of 9 of the
complete sequence.
ACCESSION AE014848 AE014188
VERSION AE014848.1 GI:23496770
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 250707)
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
AUTHORS Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
Anguoli,S., Pertea,M., Allen,J., Selegut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
MEDLINE 22255705
PUBMED 12368864
REFERENCE 2 (bases 1 to 250707)
AUTHORS Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
DIRECT SUBMISSION
TITLE Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
JOURNAL
REFERENCE 3 (bases 1 to 250707)
AUTHORS Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
DIRECT SUBMISSION
TITLE Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
JOURNAL
FEATURES
source
1..250707
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="12"
repeat_region 108..136
/rpt_type=tandem
repeat_region 372..405
/rpt_type=tandem
repeat_region 375..405
/rpt_type=tandem
repeat_region 477..541
/rpt_type=tandem
repeat_region 671..726
/rpt_type=tandem
repeat_region 1146..1171
/rpt_type=tandem
repeat_region 1522..1551
/rpt_type=tandem
repeat_region 1547..1578
/rpt_type=tandem

```

```

Best Local Similarity 70.8%; Pred. No. 2.7;
Matches 63; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 51 TTTCCTCTCTTATATCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTT 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17984 TTATTTTTCCTGTCACAGTTTAAATTCAAATATTTCTTCAGCAAAATTTCTTT 17925
QY 111 GCTAATCAATAAATATCAAAAAGAAAA 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17924 GAAATGAATTAATATGATGGAAGAAA 17896

RESULT 43
AC016751/c
LOCUS Homo sapiens BAC clone RP11-504020 from 2, complete sequence.
DEFINITION AC016751
ACCESSION AC016751
VERSION AC016751.7 GI:11415221
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 176379)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 176379)
AUTHORS Kruchowski, S., Maupin, R. and Bolla, S.
TITLE The sequence of Homo sapiens BAC clone RP11-504020
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 176379)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 176379)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 176379)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 6 (bases 1 to 176379)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 7 (bases 1 to 176379)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 29, 2000 this sequence version replaced gi:9799860.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0504020
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-514D19; the clone sequenced to the right is RP11-538A12. Actual start of this clone is at base position 1 of RP11-504020; actual end is at base position 176379 of RP11-504020.

FEATURES	Location/Qualifiers
Source	1..176379
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-504020"
	/clone_lib="RPCI-11"
	613..716
repeat_region	/rpt_family="L1"
repeat_region	722..1042
repeat_region	/rpt_family="MER2_type"
repeat_region	1215..2157
repeat_region	/rpt_family="L1"
repeat_region	2158..2449
repeat_region	/rpt_family="Alu"
repeat_region	2450..2728
repeat_region	/rpt_family="L1"
repeat_region	2904..3102
repeat_region	/rpt_family="L1"
repeat_region	3937..4069
repeat_region	/rpt_family="Alu"
repeat_region	4686..4879
repeat_region	/rpt_family="MaLR"
repeat_region	5224..5561
repeat_region	/rpt_family="MaLR"
repeat_region	5699..6006
repeat_region	/rpt_family="Alu"
repeat_region	7047..7354
repeat_region	/rpt_family="Alu"
repeat_region	7506..7647
repeat_region	/rpt_family="L1"
repeat_region	7674..8090
repeat_region	/rpt_family="L2"
repeat_region	8186..8256
repeat_region	/rpt_family="L2"
repeat_region	8262..8431
repeat_region	/rpt_family="MER2_type"
repeat_region	8470..8545
repeat_region	/rpt_family="MER2_type"

THIS PAGE BLANK (USPTO)

```
/db_xref="GOA:O97318"
/db_xref="SPTREMBL:O97318"
/translation="MGLFDKIRNIEKLNKAEALNKNNDSSMHHDOYRDDSSYIYIGNLD
LERPLVDHILNRYLPKPKYKLDADKNEYKPTGAGGGIGVNVVSESEILAVDNFNGYKL
NKSNEKKLLDDELWALNFEKSIKDIISPIGHDEKSRHNEGKEEDEDDEDD
DEDDDDSDSDYIKYRHKRKRKSUTTKYDKKEKHKRKSODHRDKRRRHRSHREK
EKDKSKRHRHRSYDYKSSRSRSTSSSTR"
Join(7801..7928,8084..8234,8392..8583)
/gene="WAL3P7.3"
/note="synonym: PFC0870w"
Join(7801..7928,8084..8234,8392..8583)
/gene="WAL3P7.3"
/note="Similarity to P. falciparum elongation factor
1-beta, (AAF27524) (276 aa), fasta scores: opt: 366, E(
1.1e-15, 53.2% identity in 141 aa overlap, revised:
shortened exon 2, splicing confirmed in P. knowlesi; Pfam:
match to PF00736 f1BD, EF-1 guanine nucleotide exchange
domain Score 61.10, E-value 7.9e-22"
/codon_start=1
/product="elongation factor 1 (EF-1); putative"
/protein_id="CAB39068.2"
/db_xref="GI:8052274"
/db_xref="GOA:O97319"
/db_xref="SPTREMBL:O97319"
/translation="WANTYDELYPLSYIILQNEGNTSKIDQANTKKPKKEVINKS
SLIIDKPYGENTDDEVLKLVKNTMEGLTWGAHKKTPEAFGLFKLQVSCVIVDDL
VNTDELITENLGLNELOKQKQMDDEYDEDDDEIGGLVOSAEIISFNKL"
Join(10596..10737,11044..11146,11373..20416,20578..20690)
/gene="WAL3P7.4"
/note="synonym: PFC0875w"
Join(10596..10737,11044..11146,11373..20416,20578..20690)
/gene="WAL3P7.4"
/note="Revised: added 3' exon, possibly spliced at 5' end,
revised: added 2x 5' exons by similarity with P. knowlesi;
signal anchor predicted by SignalP 2.0 HMM (Signal peptide
probability 0.001, signal anchor probability 0.997) with
cleavage site probability 0.000 between residues 36 and 37
Pfam match to entry PF00005 ABC_tran, ABC transporter,
score 72.50, E-value 9e-18"
/codon_start=1
/product="transporter, putative"
/protein_id="CAB39024.3"
/db_xref="GI:23477031"
/db_xref="GOA:O97278"
/db_xref="SPTREMBL:O97278"
/translation="MKIRKYLGLKLYEKRKDFLYFFPELLIPFLLISPHLFIRNI
SDKSYLEINTSSPNKINLNDTIQYVFLSNELCSEINDEIYVNHICLTPGNELM
KDFUSYANNDLINFGLYNSSEDCLNKYAIBLKSEDFKNVNSLELNDLINTHEND
KNSLLYNLHTNIINLKSEFNINDLYKDYITENEFENIKNLKINPHILKVVYKTEFYK
NFIKDKNLKNENHYQYFNDNRKLLFFYNFKNLVETKYSGLLSVENVYVKQK
NVKPSYLPGLSPQSEGKGLYKNAPLYNKVKNKOYGNPFHYYPVYVLESDDKINE
KNERNEKNERNEKNERNEKNGKNGKNDILSTNNNVVWKGIOQSYHNNNSIN
DDQKKNNNNYIYINPDGLITNVKIKRVGYALLNSNEKIFNDINILNKQFI
NFNTVDDLSEFNYFNEWYFFVLEYQFNFILNYNADILKKNLLTPRYDEINEL
NKEBETMEYKTKDEKYQVNNYINENGVDEKGYKTELKNDFFIYKMPKMK
INAPTEKNIIFRVVFLVCFLFVNICFDINKERKINENFLCCKINKYVYYSWL
LFYFILPFYNIIFTLVVVYVYKLMNYFILFYLVEFMNSLFTILCMQFSNNS
INYYATFLLPFLSSFLIIHSGASNLTFVFLIIPHSSFCLSLDFILIKNIKID
YKQIFPENISLMHLIICSIVSFLLICILNIYIYKRNKNNMLNTRYGSKSITN
DKRKSQOQTEYGINYDMSEKYPHQKDNNDPMLGSDDEDTCTYELDTLHKQV
NSOMCLRLEKIRNDRRNSNGHPDIPMLPLRNNNNNNNNNNNNNNNNNNNN
NOHRNOCYLIQVNVKPYGKHAKDQVSLTRSNRIFFVLGGENSGKSTLINTKM
ISKDSGEINFPKSYKKNKDIYMDVKNRNGLCGRICSGNGYIKFVMNRIKPCNC
DTNYDNNNNNNNDHNDHNSNGINSFNNNNVHKNKNDLEISYCSQNVIL
YENITFYETIKIFLYYNKVDKYNKRTUKIMNDLQVLYLNDLKNLMDVKKKI
SIFCLVKQDIYILDEPFIALDIKTTLKFKFPDKIKNKIIFICTHIEANNFAN
DIAVKSQIIFKSGKNSFOKLEYKFTLANIQFNGLPNEQTNYSQEKIMSALNKG
VIGSPRVANNYVDGVNRISSHSSINDKNDNNNDNNNDNNKNSVDHIDLHFD
ILVTNINEFTNRDIKKNIIKFTGTRENKNCIFPENENHYCTYKINELESKEL
LCVLNKPKNILTYOLKTIDITYIYIYITLNEKKKLLAKIQDDKIDKYLIEIDPLFLP
FQNFYFNLNLLMKNHMQPISYNFAYLRDILVRGENKNNIINSTNHDNDPYDN
MYNNHNYGKYKNNNNNSFYNNRMDDELTDIEENISSKKKKKFKIEKFSYNTYRN
```

```
KNNDENDEYGYDDNYDDNMIMSNSTNIMKKKNIKKQNIKFTCINFFTTYIKPTLLKL

Query Match
Best Local Similarity 9.0%; Score 47.4; DB 3; Length 253305;
Matches 96; Conservative 1; Mismatches 82; Indels 0; Gaps 0;

QY 38 ATTATGGGTGATTTTGCTTCCTTCTTTATACCTTTATTTATCCCAAATTTTCTTAAG 97
Db 82760 ATATATATATATATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTT 82819

QY 98 CAATATTTCTTCTGCTAATCAATAATATCAAAAGAAAAAAGTGAAGCAACGCTT 157
Db 82820 TATTTTCTTCTTATGAAAAAATAATGAAGCCCAAGGAAAAACATAATGAAATATA 82879

QY 158 GAAAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAGTTGTAAATATCTACTGT 216
Db 82880 AAATATAAATATATAAATAAATAAATAATTCATATATATAAATAAATAAATAAATA 82938

Search completed: February 14, 2004, 16:11:49
Job time : 2730 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 14:16:25 ; Search time 2621 Seconds
(without alignments)
4905.406 Million cell updates/sec

Title: US-10-027-632-1

Perfect score: 529

Sequence: 1 catggatgctgtagtctg.....tcttgagaggctcttgag 529

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_ges_hum:*

18: em_ges_inv:*

19: em_ges_pln:*

20: em_ges_vrt:*

21: em_ges_fun:*

22: em_ges_nam:*

23: em_ges_pro:*

24: em_ges_rod:*

25: em_ges_phg:*

26: em_ges_phg:*

27: em_ges_vrl:*

28: gb_ges1:*

29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	481.6	91.0	515	28	B56032 CIT-HSP-200
C 2	434.2	82.1	564	28	AQ790276 HS_2240_B
C 3	368.6	69.7	395	28	AQ803995 HS_2005_A
4	60	11.3	1091	13	BX424950 BX424950

5	57	10.8	936	13	BX404443
c	55.6	10.5	770	9	AL514205
6	55	10.4	252	13	BX425020
7	54.6	10.3	894	28	BH137168
8	54.2	10.2	427	13	BX403499
c	53.8	10.2	1201	13	BX446296
c	53.4	10.1	506	9	AL514791
c	53.4	10.1	1162	9	AL513719
c	53	10.0	234	13	BX425129
c	52.8	10.0	1104	9	AL513809
c	52.6	9.9	252	13	BX425020
c	52.4	9.9	1201	9	AL514421
c	52.4	9.9	1203	29	CNS015WU
c	52.2	9.9	699	13	BX424825
c	52.2	9.9	1201	13	BX446296
c	52.2	9.8	934	13	BX425394
c	52	9.8	923	13	BX425397
c	51.8	9.8	964	14	CB756565
c	51.6	9.8	512	13	BX424931
c	51.4	9.7	589	29	CNS03012
c	51.2	9.7	738	9	AL513907
c	51.2	9.7	838	10	BF264841
c	51.2	9.7	856	12	BM15165
c	50.8	9.6	927	13	BX371218
c	50.6	9.6	234	13	BX425129
c	50.6	9.6	381	13	BX425197
c	50.6	9.6	894	9	AL543011
c	50.4	9.5	332	13	BX446492
c	50.4	9.5	381	13	BX425197
c	50.4	9.5	940	13	BX404361
c	50.4	9.5	1201	9	AL513975
c	50.2	9.5	375	13	BX404101
c	50.2	9.5	661	9	AL514015
c	50.2	9.5	1024	13	BX349610
c	50.2	9.5	1126	13	BX446388
c	50.2	9.5	1162	9	AL513719
c	50.2	9.5	1189	9	AL513823
c	50.2	9.5	1201	9	AL513597
c	50	9.5	928	29	CNS00DKY
c	49.8	9.4	294	13	BX446415
c	49.8	9.4	314	13	BX446274
c	49.8	9.4	417	13	BX465145
c	49.8	9.4	484	14	CA722141
c	49.8	9.4	505	9	AL513553
c	49.8	9.4	535	13	BX446937
c	49.8	9.4	633	13	BX403519
c	49.8	9.4	738	9	AL513907
c	49.8	9.4	865	9	AL515307
c	49.8	9.4	1063	13	BX414736
c	49.6	9.4	358	12	BM514261
c	49.6	9.4	513	13	BQ264913
c	49.6	9.4	663	12	BJ441961
c	49.6	9.4	665	12	BQ423198
c	49.6	9.4	1009	9	AL514865
c	49.6	9.4	1179	13	BX425126
c	49.4	9.3	356	13	BX446758
c	49.4	9.3	886	9	AL514691
c	49.4	9.3	909	9	AL514871
c	49.4	9.3	1123	9	AL513999
c	49.2	9.3	422	12	B1742580
c	49.2	9.3	422	9	AW813554
c	49.2	9.3	1126	13	BX446388
c	49.2	9.3	1193	9	AL513783
c	49.2	9.3	1201	9	AL514529
c	49	9.3	197	14	CA746319
c	49	9.3	281	12	BM521967
c	49	9.3	297	13	BX425620
c	49	9.3	409	13	BX425137
c	49	9.3	425	28	AA550654
c	49	9.3	602	13	BX404150
c	49	9.3	959	13	BX403428
c	49	9.3	1056	13	BX415058
c	49	9.3	1065	9	AL515373

```

78 49 9.3 1104 9 AL513809
79 49 9.3 1189 12 BM547682
80 49 9.3 1201 9 AL551891
81 49 9.3 1998 11 BC025253
82 48.8 9.2 557 13 BQ390979
83 48.8 9.2 866 9 AL515321
84 48.8 9.2 924 9 AL515235
85 48.8 9.2 1056 13 BX415058
86 48.8 9.2 1136 9 AL514047
87 48.8 9.2 1201 29 CNS016CN
88 48.6 9.2 1279 13 BX425800
89 48.6 9.2 289 13 BX425107
90 48.6 9.2 320 13 BX446411
91 48.6 9.2 329 12 BM063269
92 48.6 9.2 386 13 BX446397
93 48.6 9.2 399 9 AL514627
94 48.6 9.2 561 13 BQ395163
95 48.6 9.2 613 28 AQ922590
96 48.6 9.2 888 28 AZ549422
97 48.6 9.2 898 13 BX432881
98 48.6 9.2 899 13 BX453223
99 48.6 9.2 909 9 AL514871
100 48.6 9.2 994 13 BX414650

AL513809 AL513809
BM547682 AGENCOURT
AL551891 AL551891
BC025253 Homo sapi
BQ390979 NISC.mq16
AL515321 AL515321
AL515235 AL515235
BX415058 BX415058
AL514047 AL514047
AL106577 Drosophil
BX425800 BX425800
BX425107 BX425107
BX446411 BX446411
BM063269 KS01053E0
BX446397 BX446397
AL514627 AL514627
BQ395163 NISC.rg13
AQ922590 RPCI-23-2
AZ549422 ENTDP49TF
BX432881 BX432881
BX453223 BX453223
AL514871 AL514871
BX414650 BX414650

ALIGNMENTS

B56032 515 bp DNA linear GSS 20-JUN-1998
CIT-HSP-2005M18.TR CIT-HSP Homo sapiens genomic clone 2005M18,
genomic survey sequence.
B56032.1 GI:2610366
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 515)
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden
,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M.
and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished
Other_GSSs: CIT-HSP-2005M18.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadam@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers
1..515
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GBS:7039412"
/db_xref="taxon:9606"
/clone="2005M18"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/notes="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
152 a 102 c 112 g 149 t

RESULT 1
B56032/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
AQ790276 564 bp DNA linear GSS 03-AUG-1999
HS_2240_B1_F01_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=2240 Col=1 Row=L, genomic survey
sequence.
AQ790276
AQ790276.1 GI:5697900
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
```


Invitrogen. This sequence belongs to sequence cluster 6437.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CL0BB0062G12FP1&cluster=6437.r>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CL0BB0062G12FP1.

FEATURES

source

1..770
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CL0BB0062G12"
 /tissue_type="NEUROBLASTOMA"
 /clone_lib="Homo sapiens NEUROBLASTOMA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
 Location/Qualifiers
 1..770
 BASE COUNT 216 a 35 c 193 g 181 t 145 others
 ORIGIN

Query Match 10.5%; Score 55.6; DB 9; Length 770;
 Best Local Similarity 52.7%; Pred. No. 1.5e+02;
 Matches 77; Conservative 17; Mismatches 52; Indels 0; Gaps 0;

Qy 24 TGATTATATGCGATTATGCGGTGATTTGCTTCCTTTATATCTTTATTTATTCCTCC 83
 Db 277 KTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 218
 Qy 84 AAATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACAAC 143
 Db 217 TTTTCTTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACAAC 158
 Qy 144 TGAAGCAACGCTTGAAAAAGGAAA 169
 Db 157 AAAAAAAGAAAAAAGAAAAAAGAAAAA 132

RESULT 7
 BX425020 252 bp mRNA linear EST 15-MAY-2003
 DEFINITION BX425020 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA007ZD06
 3-PRIME, mRNA sequence.

ACCESSION BX425020
 VERSION BX425020
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 252)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France

Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8225.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CL0BA007ZD06FP1&cluster=8225.r>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CL0BA007ZD06FP1.
 Location/Qualifiers
 1..252

FEATURES

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

/clone="CL0BA007ZD06"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
 Location/Qualifiers
 84 a 20 c 24 g 89 t 35 others
 BASE COUNT
 ORIGIN

Query Match 10.4%; Score 55; DB 13; Length 252;
 Best Local Similarity 47.9%; Pred. No. 3.2e+02;
 Matches 79; Conservative 23; Mismatches 63; Indels 0; Gaps 0;

Qy 5 GTGATGCTGTAGTCGTGCTGATTTATATGCTGATTTATGCTGATTTGCTGCTCTCTT 64
 Db 68 GGGATACCTAGACGGCGCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 127
 Qy 65 TATATCTTTTATTTATTTCCAAATTTTCTTAAGCAATATTTCTTTGCTAATCAATAAT 124
 Db 128 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 187
 Qy 125 TATCAAAAGAAAAAAGAAAAAAGCAACGCTTGAAAAAAGGAAA 169
 Db 188 WAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 232

RESULT 8
 BH137168 894 bp DNA linear GSS 07-AUG-2001
 LOCUS ENTOKI4TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION genomic, genomic survey sequence.

ACCESSION BH137168
 VERSION BH137168.1 GI:15096229
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica
 ORGANISM Entamoeba histolytica
 Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.

REFERENCE 1 (bases 1 to 894)
 AUTHORS Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
 TITLE Determination of clone end sequences from Entamoeba histolytica
 JOURNAL HMI:IMSS sheared DNA library (2001)
 COMMENT Unpublished
 CONTACT: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library
 Seq primer: M13-Reverse
 Class: shotgun
 High quality sequence start: 24
 High quality sequence stop: 522.
 Location/Qualifiers
 1..894
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith,

FEATURES

source

/organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith,


```

RESULT 11
AL514791
LOCUS
DEFINITION
AL514791 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB015ZE10 3-PRIME, mRNA sequence.
ACCESSION
AL514791.2 GI:30464676
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 506)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12778284.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1606.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBB015ZE10FPl&cluster=1606.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CLOBB015ZE10FPl.
Location/Qualifiers
1..506
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB015ZE10"
/tissue_type="NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 332 a 24 c 24 g 102 t 24 others
ORIGIN
Query Match 10.1%; Score 53.4; DB 9; Length 506;
Best Local Similarity 54.5%; Pred. NO. 3.7e+02;
Matches 78; Conservative 12; Mismatches 53; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGGGTGATTTGCTTCTTCTTATCTTTTATTTATTTATTTCCCAA 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TTTTCTTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 168
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 87 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 228
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 147 AAGCAACGCTTGAAAAAGGAAA 169
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 AAAAAAAGAAAAAAGAAAAA 251
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AL513719/c
LOCUS
DEFINITION
AL513719 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBB007ZH01
3-PRIME, mRNA sequence.
ACCESSION
AL513719
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1162)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12777213.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 817.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBB007ZH01FPl&cluster=817.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CLOBB007ZH01FPl.
Location/Qualifiers
1..1162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB007ZH01"
/tissue_type="PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 394 a 229 c 143 g 98 t 298 others
ORIGIN
Query Match 10.1%; Score 53.4; DB 9; Length 1162;
Best Local Similarity 39.9%; Pred. NO. 2.4e+02;
Matches 57; Conservative 40; Mismatches 46; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGGGTGATTTGCTTCTTCTTATCTTTTATTTATTTCCCAA 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251 TTTTCTTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 192
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 87 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 147 AAGCAACGCTTGAAAAAGGAAA 169
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 AAAAAAAGAAAAAAGAAAAA 109
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
BX425129/c
LOCUS
DEFINITION
BX425129 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBB011ZH06
3-PRIME, mRNA sequence.
ACCESSION
BX425129
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 234)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9703.f For

```

```

REFERENCE
AL514791
LOCUS
DEFINITION
AL514791 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB015ZE10 3-PRIME, mRNA sequence.
ACCESSION
AL514791.2 GI:30464676
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 506)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12778284.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1606.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBB015ZE10FPl&cluster=1606.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CLOBB015ZE10FPl.
Location/Qualifiers
1..506
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB015ZE10"
/tissue_type="NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 332 a 24 c 24 g 102 t 24 others
ORIGIN
Query Match 10.1%; Score 53.4; DB 9; Length 506;
Best Local Similarity 54.5%; Pred. NO. 3.7e+02;
Matches 78; Conservative 12; Mismatches 53; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGGGTGATTTGCTTCTTCTTATCTTTTATTTATTTCCCAA 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TTTTCTTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 168
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 87 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 228
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 147 AAGCAACGCTTGAAAAAGGAAA 169
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 AAAAAAAGAAAAAAGAAAAA 251
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AL513719/c
LOCUS
DEFINITION
AL513719 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBB007ZH01
3-PRIME, mRNA sequence.
ACCESSION
AL513719
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1162)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12777213.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 817.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBB007ZH01FPl&cluster=817.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CLOBB007ZH01FPl.
Location/Qualifiers
1..1162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB007ZH01"
/tissue_type="PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 394 a 229 c 143 g 98 t 298 others
ORIGIN
Query Match 10.1%; Score 53.4; DB 9; Length 1162;
Best Local Similarity 39.9%; Pred. NO. 2.4e+02;
Matches 57; Conservative 40; Mismatches 46; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGGGTGATTTGCTTCTTCTTATCTTTTATTTATTTCCCAA 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251 TTTTCTTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 192
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 87 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 TTTTCTTAAAGCAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 147 AAGCAACGCTTGAAAAAGGAAA 169
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 AAAAAAAGAAAAAAGAAAAA 109
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
BX425129/c
LOCUS
DEFINITION
BX425129 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBB011ZH06
3-PRIME, mRNA sequence.
ACCESSION
BX425129
VERSION
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 234)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9703.f For

```

[illegible]

```
Best Local Similarity 49.0%; Pred. No. 6.8e+02;
Matches 70; Conservative 22; Mismatches 51; Indels 0; Gaps 0;

QY 27 TTTTATCTGATTTATGGGTGATTTGCTTCCTCTTTATCTTTATTTATTTCCCAAA 86
Db 245 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 186

QY 87 TTTTCTTAAGCAATATTTCTTCTGCTAATCAATAAATTTCAAAAGAAAAAACTGA 146
Db 185 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 126

QY 147 AAGCACTCTGAAAAAGGAAA 169
Db 125 AAAAAAAGAAAAAAGAAAAA 103

RESULT 16
AL514421
LOCUS 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION AL514421 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
ACCESSION CLOBB0102G01 3-PRIME, mRNA sequence.
VERSION AL514421
KEYWORDS AL514421.2 GI:30464306
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1201)
JOURNAL Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
COMMENT Full-length cDNA libraries and normalization
On Feb 13, 2001 this sequence version replaced gi:12777915.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7847.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBB0102G01P1&cluster=7847.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBB0102G01FPI.
FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB0102G01"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector..
Library was not normalized."
BASE COUNT 249 a 93 c 32 g 650 t 177 others
ORIGIN
Query Match 9.9%; Score 52.4; DB 9; Length 1201;
Best Local Similarity 46.2%; Pred. No. 3.3e+02;
Matches 98; Conservative 12; Mismatches 102; Indels 0; Gaps 0;

QY 25 GATTATATGCTGATTTATGGGTGATTTGCTTCCTCTTTATCTTTATTTATTTCCCA 84
Db 546 GNTTNTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTAATATNTTTTTTTTTTTTTTT 605

QY 85 AATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATTTCAAAAGAAAAAACT 144
Db 606 NTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 665
```

```
QY 145 GAAAGCAACGCTTGAAAAAGAAAGTTAGCCCTATCGGGTATATTTTGGAGTTGTAAA 204
Db 666 AWAAGAAAAATTTTATTTAAAAAATTTTAAATAATTTTCTTTTAAATATAWA 725

QY 205 ATACTACRTGTTCTCTTCTTAAGTCCCACTCT 236
Db 726 WMAAAAAAATTTCTTMTMTHTATATATCCCCCTT 757

RESULT 17
CNS015WU 1203 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence 896 end of BAC
DEFINITION BACN15E10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL106008
VERSION AL106008.1 GI:5619558
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1203)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
source
1..1203
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15E10"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/notes="end : SP6"
BASE COUNT 288 a 103 c 235 g 253 t 324 others
ORIGIN
Query Match 9.9%; Score 52.4; DB 29; Length 1203;
Best Local Similarity 30.9%; Pred. No. 3.3e+02;
Matches 102; Conservative 73; Mismatches 155; Indels 0; Gaps 0;

QY 26 ATTATATGCTGATTTATGGGTGATTTGCTTCCTCTTTATCTTTATTTATTTCCCAA 85
Db 1154 AAWAAWAAWTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1095

QY 86 ATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATTTCAAAAGAAAAAACTG 145
Db 1094 AAWAAWAAWATYTTTWTWAAWATWAAWAAWAAWAAWAAWAAWAAWAAWAAW 1035

QY 146 AAAGCAACGCTTGAAAAAGAAAGTTAGCCCTATCGGGTATATTTTGGAGTTGTAAA 205
Db 1034 AAWTWTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 975

QY 206 TACTACRTGTTCTCTTCTTAAGTCCCACTCTCTGTTTCTTTGACGAGAAAGAAAGC 265
Db 974 AWAATWATATWATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 915

QY 266 AGTCACTCTCTTCCATGACAAACCCATGACGGCTGCTGCTGCTCATCACAGG 325
Db 914 TAWCAMYCCCCYYTTTMYTWTAAAAACCCMCCYCYMYTWCACCCCTTTCCTTTC 855
```



```
BASE COUNT      330 a      280 c      224 g      129 t      1 others
ORIGIN

Query Match      9.8%; Score 51.6; DB 14; Length 964;
Best Local Similarity 56.5%; Pred. No. 4.7e+02;
Matches 96; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 12 TGTAGTCGGCTGATTTATATGCTGATTTATGGGATTTTGTCTTCCTTCTTTATACCT 71
DB 31 TTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 90
QY 72 TTATTATTCCTCAATTTTCTTAAGCAATATTTCTTCTGAATCAATPAATATCAAA 131
DB 91 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 150
QY 132 AGAAAAAACTGAAGCAACGCTGAAAAAGGAAAGTATGCCCTATC 181
DB 151 AAAAAAAACCCAAACCAAGGGGGGAAACACGGAAGGAGGCAACC 200

RESULT 23
BX424931/c
LOCUS
DEFINITION      512 bp mRNA linear EST 15-MAY-2003
ACCESSION      BX424931
VERSION        BX424931.1 GI:30782418
KEYWORDS
SOURCE
ORGANISM        Homo sapiens (human)
REFERENCE
AUTHORS        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
COMMENT        1 (bases 1 to 512)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1009.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BA004ZC11P1&cluster=1009.f. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BA004ZC11P1.
Location/Qualifiers
1. .512
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA004ZC11"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT      143 a      143 c      50 g      111 t      65 others
ORIGIN

Query Match      9.7%; Score 51.4; DB 13; Length 512;
Best Local Similarity 49.1%; Pred. No. 6.9e+02;
Matches 82; Conservative 17; Mismatches 68; Indels 0; Gaps 0;

QY 3 TGTGATGCTAGTCGCTGATTTATATGCTGATTTATGGTGATTTTGTCTTCCTTC 62
DB 296 KKKKKKKKTTTTTKGGGGKKTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 237
QY 63 TTTATACTTTTATTTATCCCAATTTTCTTTAAGCAATATTTCTTCTGCTAATCAATA 122

BASE COUNT      330 a      280 c      224 g      129 t      1 others
ORIGIN

Query Match      9.8%; Score 51.6; DB 14; Length 964;
Best Local Similarity 56.5%; Pred. No. 4.7e+02;
Matches 96; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 12 TGTAGTCGGCTGATTTATATGCTGATTTATGGGATTTTGTCTTCCTTCTTTATACCT 71
DB 31 TTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 90
QY 72 TTATTATTCCTCAATTTTCTTAAGCAATATTTCTTCTGAATCAATPAATATCAAA 131
DB 91 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 150
QY 132 AGAAAAAACTGAAGCAACGCTGAAAAAGGAAAGTATGCCCTATC 181
DB 151 AAAAAAAACCCAAACCAAGGGGGGAAACACGGAAGGAGGCAACC 200

RESULT 24
BX424931/c
LOCUS
DEFINITION      589 bp DNA linear GSS 01-SEP-2000
ACCESSION      AL253412
VERSION        AL253412.1 GI:7974424
KEYWORDS        GSS; genome survey sequence.
SOURCE          Tetraodon nigroviridis
ORGANISM        Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1
Roest Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
PUBMED
10835645

Roest Crollius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
PUBMED
10899143

Direct Submision
Genoscope.
3 (bases 1 to 589)
Direct Submision
Genoscope.
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1. .589
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="043121"
/clone_lib="G"
/notes="Genoscope sequence ID : COBG043AE11LP1-end : T7"
BASE COUNT      378 a      20 c      38 g      98 t      55 others
ORIGIN

Query Match      9.7%; Score 51.2; DB 29; Length 589;
Best Local Similarity 48.2%; Pred. No. 6.8e+02;
Matches 92; Conservative 11; Mismatches 88; Indels 0; Gaps 0;

QY 2 ATGGTGATGCTAGTCGCTGATTTATATGCTGATTTATGGTGATTTTGTCTTCCTT 61
DB 470 ATTTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 411
QY 62 CTTTATACCTTTTATTTATTTCCCAATTTTCTTTAAGCAATATTTCTTGTCTAATCAATA 121
DB 410 TTWTATTTTNTATANNATCTANATCTTTTAAATTTTATTAATTAATTAATTAAT 351
```

```

Oy 122 AATTATCAAGAAAAAAGTCTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCTATC 181
    ||| ||| :|||: :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 AATRAAGAAAAAARAWAAAAAAMMMMAAAAAAATAAATAAAWANNNTNNAWT 291
    ||| ||| :|||: :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 182 GGGTATATTTT 192
    ||| ||| :|||: :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 GTTTTTTTTTT 280
    ||| ||| :|||: :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 25
AL513907 738 bp mRNA linear EST 08-MAY-2003
LOCUS AL513907 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA006ZC11
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION AL513907
VERSION AL513907.2 GI:30463792
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 738)
Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12777401.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9836.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgt-bin/Cluster.cgi?seq=CLOBA006ZC11&plcluster=9836.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA006ZC11FPI.
FEATURES
    source
    1..738
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CLOBA006ZC11"
    /tissue_type="PLACENTA"
    /clone_lib="Homo sapiens PLACENTA"
    /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
    with a NotI-oligo(dT) primer. Five prime end enriched,
    double-strand cDNA was digested with Not I and cloned into
    the Not I and EcoRV sites of the pCMVSPORT 6 vector.
    Library was not normalized."
    BASE COUNT 184 a 28 c 95 g 342 t 89 others
ORIGIN
    Query Match 9.7%; Score 51.2; DB 9; Length 738;
    Best Local Similarity 52.0%; Pred. No. 6.1e+02;
    Matches 77; Conservative 14; Mismatches 57; Indels 0; Gaps 0;

Oy 27 TTATATGCTGATTATGGGTAATTTGCTTCCTTCTTTATCTTTTATTTATCCCAAA 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 87 TTTTCTTAAGCAATATTTCTTCTGCTATCATTAATATCAAGAAAAAAGAACTGA 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 TTTTCTTTTWWAAATTTTATTTTAAAAAATAAATAAATAAATAAATAAATAAATAA 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 147 AAGCAACGCTTGAAAAAGGAAAGTTAG 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 26

```

```

BF264841/c
LOCUS BF264841 838 bp mRNA linear EST 23-OCT-2001
DEFINITION HV_CEA0010112f Hordeum vulgare seedling green leaf EST library
HVCDA0004 (Blumeria challenged) Hordeum vulgare subsp. vulgare
cDNA clone HV_CEA0010112f, mRNA sequence.
ACCESSION BF264841
VERSION BF264841.1 GI:11195835
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
1 (bases 1 to 838)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D.,
Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi
,D.W., Fenton,R.D., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Blumeria infected incompatible (Mla13)
seedling leaf cDNA library
Unpublished
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 365
Seq primer: AATTAAACCTCCTCACTAAAGGG
High quality sequence stop: 813.
FEATURES
    source
    1..838
    /organism="Hordeum vulgare subsp. vulgare"
    /mol_type="mRNA"
    /cultivar="C116155 (Mla13)"
    /db_xref="taxon:112509"
    /clone="HV_CEA0010112f"
    /tissue_type="seedling green leaf"
    /lab_host="TJC121"
    /clone_lib="Hordeum vulgare seedling green leaf EST
    library HVCDA0004 (Blumeria challenged)"
    /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
    C.I. 16155 (Mla13) plants were greenhouse grown in the R
    Wise lab at Iowa State University, Ames, IA; 7 day old
    green seedlings were challenged with isolate A27 (AvrMla13
    ) of Blumeria graminis f. sp. hordei, and leaves were
    harvested 20 and 24 hr post-inoculation and snap frozen;
    uninoculated leaves were harvested 20 hr post-inoculation
    (Wei, Wise). In the TJ Close lab at the University of
    California, Riverside, total RNA was prepared from each
    sample pool, equal quantities of all three RNA pools were
    combined, poly(A) RNA was purified from the mixture, one
    cDNA library was made, and 1 million pfu were in vivo
    excised to give pBluescript SK(-) cDNA phagemids (Choi,
    Close). Phagemids were plated and picked at the Clemson
    University Genomics Institute (CUGI) (Begum, Palmer,
    Frisch, Atkins and Wing). plasmid DNA preparations, DNA
    sequencing and sequence analysis were performed at CUGI
    (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
    The sequence has been trimmed to remove vector sequence
    and contains a minimum of 100 bases of phred value 20 or
    above. For more details on library preparation and
    sequence analysis see
    http://www.genome.clemson.edu/projects/barley. To order
    this clone see http://www.genome.clemson.edu/orders
    see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
    Genetically and physically anchored EST resources for
    barley genomics. Barley Genetics Newsletter 31:29-30.
    (http://wheat.pw.usda.gov/gspages/bgn/31/cover.html)"
    BASE COUNT 296 a 163 c 102 g 275 t 2 others
ORIGIN

```


VERSION	EST.
BA423197.1	Homo sapiens (human)
GI:30769941	

Db

236 TTHHWITWWAHAAAAWTTTTTYYAWWWWMAAAAWWWHHHWWWWAAAAA 177

LOCUS	AL543011	894 bp	mRNA	linear	EST 12-MAY-2003

AL543011	894 bp	mRNA	linear	EST 12-MAY-2003

```

DEFINITION AL543011 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YL22
5-PRIME, mRNA sequence.
ACCESSION AL543011
VERSION AL543011.2 GI:30548727
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 894)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12875489.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9876.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE013DF11QPl&cluster=9876.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE013DF11QPl.

FEATURES
source
1..894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YL22"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 234 a 158 c 203 g 237 t 62 others
ORIGIN
Query Match 9.6%; Score 50.6; DB 9; Length 894;
Best Local Similarity 41.5%; Pred. No. 6.7e+02;
Matches 81; Conservative 33; Mismatches 81; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTATGAGTATTTGCTTCCTCTTATACCTTTTATTTATTTATCCCAA 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D691 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 750
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 87 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 146
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D751 TTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 810
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 147 AAGCAACGCTTGAAAGGAAAGTGTAGCCCTATCGGCTATATTTTGGAGTTGTAAAT 206
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D811 AAAAAAAAAAAAAAARADGKAABAAATAAATAAATAAATAAATAAATAAATAAATAA 870
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 207 ACTACTGTGTCCTT 221
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D871 ATTAAKAWTTTTTT 885
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 32
BX446492/c
DEFINITION BX446492 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA009ZC08
3-PRIME, mRNA sequence.
ACCESSION BX446492
VERSION BX446492.1 GI:31033732
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 332)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6609.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA009ZC08FP1&cluster=6609.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA009ZC08FP1.

FEATURES
source
1..332
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA009ZC08"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 143 a 25 c 24 g 122 t 18 others
ORIGIN
Query Match 9.5%; Score 50.4; DB 13; Length 332;
Best Local Similarity 53.2%; Pred. No. 1.2e+03;
Matches 84; Conservative 9; Mismatches 65; Indels 0; Gaps 0;

QY 12 TGTACTCGTGGCTGATTATATGCTGATTATGCTGATTATGCTGCTTCTCTTATACCT 71
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D292 TTTKKTITTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 233
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 72 TTATTTATTTCCAAATTTTCTTAAGCAAAATTTTCTTGTCTTAATCAATAAATATCAAA 131
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D232 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 173
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 132 AGAAAAAAACCTGAAGCAAGCTTTGAAAAAGGAAA 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D172 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 135
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 33
BX425197
DEFINITION BX425197 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB003ZH07 3-PRIME, mRNA sequence.
ACCESSION BX425197
VERSION BX425197.1 GI:30769941
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 381)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6437.r For

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 332)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6609.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA009ZC08FP1&cluster=6609.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA009ZC08FP1.

FEATURES
Location/Qualifiers
1..332
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA009ZC08"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 143 a 25 c 24 g 122 t 18 others
ORIGIN
Query Match 9.5%; Score 50.4; DB 13; Length 332;
Best Local Similarity 53.2%; Pred. No. 1.2e+03;
Matches 84; Conservative 9; Mismatches 65; Indels 0; Gaps 0;

QY 12 TGTACTCGTGGCTGATTATATGCTGATTATGCTGATTATGCTTCTCTTATACCT 71
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D292 TTTKKTITTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 233
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 72 TTATTTATTTCCAAATTTTCTTAAGCAAAATTTTCTTGTCTTAATCAATAAATATCAAA 131
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D232 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 173
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 132 AGAAAAAAACCTGAAGCAAGCTTTGAAAAAGGAAA 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D172 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 135
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 33
BX425197
DEFINITION BX425197 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB003ZH07 3-PRIME, mRNA sequence.
ACCESSION BX425197
VERSION BX425197.1 GI:30769941
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 381)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6437.r For

```



```
Best Local Similarity 49.3%; Pred. No. 6.1e+02;
Matches 72; Conservative 19; Mismatches 55; Indels 0; Gaps 0;

QY 24 TGATTATATGCTGATTATGGGTGATTTTGGCTTCCTTTATATCTTTTATTATCC 83
Db 98 TTTTATATGCTGATTATGGGTGATTTTGGCTTCCTTTATATCTTTTATTATCC 157
QY 84 AAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAAAC 143
Db 158 AAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAAAC 217
QY 144 TGAACCAACGCTTGAAAAAGGAAA 169
Db 218 AAAAAAAGGCTTGAAAAAGGAAA 243

RESULT 36
BX404101/c
LOCUS
DEFINITION BX404101 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
ACCESSION CL0BB015ZC10 3-PRIME, mRNA sequence.
VERSION BX404101
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 375)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7985.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BB015ZC10FP1&cluster=7985.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BB015ZC10FP1.

FEATURES
source
Location/Qualifiers
1..375
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB015ZC10"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 143 a 20 c 20 g 172 t 20 others
ORIGIN
Query Match 9.5%; Score 50.2; DB 13; Length 375;
Best Local Similarity 55.2%; Pred. No. 1.2e+03;
Matches 79; Conservative 8; Mismatches 56; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGGGTGATTTTGGCTTCCTTTATATCTTTTATTATCCAA 86
Db 326 TTTTATATGCTGATTATGGGTGATTTTGGCTTCCTTTATATCTTTTATTATCCAA 267
QY 87 TTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAAACCTGA 146
Db 266 TTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAAACCTGA 207
QY 147 AAGCAACGCTTGAAAAAGGAAA 169
```

```
Db 206 AAAAAAAGGCTTGAAAAAGGAAA 184

RESULT 37
AL514015
LOCUS
DEFINITION AL514015 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA011ZG02
ACCESSION 3-PRIME, mRNA sequence.
VERSION AL514015
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 661)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12777509.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 108.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BA011ZG02FP1&cluster=108.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BA011ZG02FP1.

FEATURES
source
Location/Qualifiers
1..661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA011ZG02"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT.6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 151 a 107 c 159 g 155 t 89 others
ORIGIN
Query Match 9.5%; Score 50.2; DB 9; Length 661;
Best Local Similarity 51.0%; Pred. No. 8.8e+02;
Matches 73; Conservative 16; Mismatches 54; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGGGTGATTTTGGCTTCCTTTATATCTTTTATTATCCAA 86
Db 123 TTTTATATGCTGATTATGGGTGATTTTGGCTTCCTTTATATCTTTTATTATCCAA 182
QY 87 TTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAAACCTGA 146
Db 183 TTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAAACCTGA 242
QY 147 AAGCAACGCTTGAAAAAGGAAA 169
Db 243 AAAAAAAGGCTTGAAAAAGGAAA 265

RESULT 38
BX349610
LOCUS
DEFINITION BX349610 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION clone CS0D1021YP19 3-PRIME, mRNA sequence.
VERSION BX349610
COMMENT BX349610.1 GI:30365363
```


VERSION	EST.	ORGANISM
BX446274.1	Homo sapiens (human)	Homo sapiens
GI:30776405		

GenCore version 5.1.6
Copyright (c) 1993 - 2004. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 13:45:13 ; Search time 325 Seconds

(without alignments)
5995.821 Million call updates/sec

Title: US-10-027-632-1

Perfect score: 529

Sequence: 1 catggtgatgtgtagtctg.....ctctggagaggctcttgag 529

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 184181637 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.6	99.9	529	13	US-10-027-632-1
2	528.6	99.9	529	14	US-10-027-632-1
3	527.6	99.7	535	13	US-10-027-632-54345
4	527.6	99.7	535	13	US-10-027-632-300106
5	527.6	99.7	535	14	US-10-027-632-54345
6	527.6	99.7	535	14	US-10-027-632-300106
7	51.4	9.7	321	13	US-09-814-353-17876
8	51	9.6	284	13	US-09-814-353-17553
9	49.6	9.4	424	10	US-09-960-352-11218
10	48.2	9.1	316	13	US-09-814-353-16036
11	47.4	9.0	113515	13	US-10-311-455-2148
12	47.2	8.9	597	13	US-09-814-353-4872
13	47.2	8.9	597	13	US-09-814-353-11169
14	46.6	8.8	312	10	US-09-960-352-8414
15	46.6	8.8	383	13	US-09-814-353-18006

46.4	8.8	480	12	US-10-131-827-8192	Sequence 8192, Ap
46.2	8.7	312	10	US-09-960-352-8414	Sequence 8414, Ap
46	8.7	3673778	13	US-10-312-841-2	Sequence 2, Appli
45.8	8.7	463	13	US-09-814-353-4849	Sequence 4849, Ap
45.8	8.7	463	13	US-09-814-353-11146	Sequence 11146, A
45.8	8.7	3673778	13	US-10-312-841-1	Sequence 1, Appli
45.6	8.6	406	13	US-09-814-353-17782	Sequence 17782, A
45	8.5	359	13	US-09-814-353-17359	Sequence 17359, A
44.8	8.5	321	13	US-09-814-353-17876	Sequence 17876, A
44.8	8.5	723	13	US-09-814-353-4726	Sequence 4726, Ap
44.8	8.5	723	13	US-09-814-353-11024	Sequence 11024, A
44.4	8.4	428	13	US-09-814-353-17588	Sequence 17588, A
44.4	8.4	425	10	US-09-834-975-451	Sequence 451, App
44.4	8.4	664	9	US-09-739-254-66	Sequence 66, Appl
44.4	8.4	664	13	US-10-055-098-66	Sequence 66, Appl
44.4	8.4	664	15	US-10-054-988-66	Sequence 66, Appl
44.4	8.4	6078	13	US-10-311-455-1110	Sequence 1110, Ap
44.2	8.4	3673778	13	US-10-312-841-1	Sequence 1, Appli
44	8.3	281	13	US-09-814-353-17383	Sequence 17383, A
44	8.3	600	15	US-10-198-846-8434	Sequence 8434, Ap
43.8	8.3	236	10	US-09-960-352-12183	Sequence 12183, A
43.8	8.3	281	13	US-09-814-353-17383	Sequence 17383, A
43.8	8.3	610	13	US-09-814-353-5380	Sequence 5380, Ap
43.8	8.3	610	13	US-09-814-353-11667	Sequence 11667, A
43.6	8.2	9884	15	US-10-172-086-5	Sequence 5, Appli
43.6	8.2	6590	13	US-10-311-455-1449	Sequence 1449, Ap
43.4	8.2	424	10	US-09-960-352-11218	Sequence 11218, A
43.2	8.2	380	10	US-09-960-352-9335	Sequence 9335, Ap
43.2	8.2	397	13	US-09-814-353-17585	Sequence 17585, Ap
43.2	8.2	430	15	US-10-078-090-87	Sequence 87, Appl
43	8.1	255	13	US-09-930-213-715	Sequence 715, App
43	8.1	512	13	US-09-814-353-5072	Sequence 5072, Ap
43	8.1	512	13	US-09-814-353-11364	Sequence 11364, A
43	8.1	180557	14	US-10-003-806-6	Sequence 6, Appli
43	8.1	180557	14	US-10-003-806-9	Sequence 9, Appli
43	8.1	368004	10	US-09-949-654-3	Sequence 3, Appli
43	8.1	3673778	13	US-10-312-841-2	Sequence 2, Appli
42.8	8.1	358	13	US-10-198-846-9603	Sequence 9603, Ap
42.8	8.1	383	13	US-09-814-353-18006	Sequence 18006, A
42.6	8.1	515	13	US-09-814-353-6142	Sequence 6142, Ap
42.6	8.1	515	13	US-09-814-353-12421	Sequence 12421, A
42.4	8.0	308	10	US-09-960-352-7670	Sequence 7670, Ap
42.4	8.0	384	13	US-09-814-353-2534	Sequence 2534, Ap
42.4	8.0	384	13	US-09-814-353-8870	Sequence 8870, Ap
42.4	8.0	524	13	US-10-125-968-1258	Sequence 1258, Ap
42.4	8.0	1081	13	US-10-027-632-118342	Sequence 118342, Ap
42.4	8.0	1081	13	US-10-027-632-118343	Sequence 118343, Ap
42.4	8.0	1081	14	US-10-027-632-118343	Sequence 118343, Ap
42.4	8.0	1081	14	US-10-027-632-118343	Sequence 118343, Ap
42.2	8.0	6731	13	US-10-311-455-1033	Sequence 1033, Ap
42.2	8.0	277	10	US-09-960-352-12673	Sequence 12673, A
42.2	8.0	302	13	US-09-814-353-5335	Sequence 5335, Ap
42.2	8.0	302	13	US-09-814-353-11622	Sequence 11622, A
42.2	8.0	380	10	US-09-960-352-9335	Sequence 9335, Ap
42.2	8.0	390	13	US-09-814-353-17808	Sequence 17808, A
42.2	8.0	418	10	US-09-960-352-4845	Sequence 4845, Ap
42.2	8.0	435	13	US-09-814-353-5103	Sequence 5103, Ap
42.2	8.0	435	13	US-09-814-353-11335	Sequence 11335, A
42.2	8.0	476	13	US-09-814-353-4700	Sequence 4700, Ap
42.2	8.0	476	13	US-09-814-353-10999	Sequence 10999, A
42.2	8.0	550	13	US-09-814-353-5098	Sequence 5098, Ap
42.2	8.0	550	13	US-09-814-353-11390	Sequence 11390, A
42.2	8.0	1066	9	US-09-770-445-141	Sequence 141, App
42.2	8.0	5359	13	US-10-311-455-1580	Sequence 1580, Ap
42.2	8.0	5359	13	US-10-240-453-146	Sequence 146, App
42.2	8.0	5359	13	US-10-240-453-257	Sequence 257, App
42.2	8.0	6533	13	US-09-814-353-4844	Sequence 4844, Ap
42	7.9	299	13	US-09-814-353-11141	Sequence 11141, A
42	7.9	299	13	US-09-814-353-4872	Sequence 4872, Ap
42	7.9	597	13	US-09-814-353-11169	Sequence 11169, A
42	7.9	597	13	US-09-814-353-11169	Sequence 11169, A
42	7.9	4163	13	US-10-311-455-2303	Sequence 2303, Ap

c 89 42 7.9 13449 13 US-10-311-455-1357 Sequence 1357, Ap
c 90 41.8 7.9 284 13 US-09-814-353-17553 Sequence 17553, A
91 41.8 7.9 396 9 US-09-825-294-33 Sequence 33, Appl
92 41.8 7.9 396 10 US-09-970-966-33 Sequence 33, Appl
93 41.8 7.9 396 12 US-10-369-186-33 Sequence 33, Appl
94 41.8 7.9 396 13 US-10-361-811-33 Sequence 33, Appl
95 41.8 7.9 396 16 US-10-212-677-33 Sequence 33, Appl
96 41.8 7.9 513509 11 US-09-754-853A-4 Sequence 4, Appl
c 97 41.6 7.9 298 10 US-09-960-352-1004 Sequence 1004, Ap
98 41.6 7.9 307 13 US-09-814-353-17818 Sequence 17818, A
99 41.6 7.9 384 13 US-09-814-353-18604 Sequence 18604, A
100 41.6 7.9 385 13 US-09-814-353-5368 Sequence 5368, Ap

ALIGNMENTS

RESULT 1
US-10-027-632-1
; Sequence 1, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-1

Query Match 99.9%; Score 528.6; DB 13; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.5e-119;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGTGTAGCTGTAGTGGCTGATTTATATGCTGATTTAATGGGTGATTTGCTTCT 60
DB 1 CATGGTGTAGCTGTAGTGGCTGATTTATATGCTGATTTAATGGGTGATTTGCTTCT 60
QY 61 TCTTTATACCTTTTATTTATCCCAATTTTCTTAAGCAAAATATTTCTTGCTAATCAAT 120
DB 61 TCTTTATACCTTTTATTTATTCCTAAATTTTCTTAAGCAAAATATTTCTTGCTAATCAAT 120
QY 121 AAATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTAT 180
DB 121 AAATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTAT 180
QY 181 CGGGTATATTTTGGAGTTGTAAATATCTACRTGTCTCTTCTTAAGTCCCACTCTCTGT 240
DB 181 CGGGTATATTTTGGAGTTGTAAATATCTACRTGTCTCTTCTTAAGTCCCACTCTCTGT 240
QY 241 TTTCTTTGAGCAGGAAGAAAGCAAGTCACTCTCTTCCATGACAAACCCATGACC 300
DB 241 TTTCTTTGAGCAGGAAGAAAGCAAGTCACTCTCTTCCATGACAAACCCATGACC 300

QY 301 GGCTGCTTGCCTGCTCCATCACAGGAAGGTGCAGCAGAGCCCTTTTCTTCCAGGTTCT 360
DB 301 GGCTGCTTGCCTGCTCCATCACAGGAAGGTGCAGCAGAGCCCTTTTCTTCCAGGTTCT 360
QY 361 GGGCACCTTTGAACTATCTCTGCTTGTAGTAAAGAACTGCCCCAAGTGACGACACATCA 420
DB 361 GGGCACCTTTGAACTATCTCTGCTTGTAGTAAAGAACTGCCCCAAGTGACGACACATCA 420
QY 421 AGCTGTTGGGACACTTCAGAAATAAACCAACCACTGTATTACTGGTCTTTTGACATCT 480
DB 421 AGCTGTTGGGACACTTCAGAAATAAACCAACCACTGTATTACTGGTCTTTTGACATCT 480
QY 481 CTCTGCTCAGGCGCAAGTCACTGCTCTCTTGGAGAGGGCTTCTGAG 529
DB 481 CTCTGCTCAGGCGCAAGTCACTGCTCTCTTGGAGAGGGCTTCTGAG 529

RESULT 2

US-10-027-632-1
; Sequence 1, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-1

Query Match 99.9%; Score 528.6; DB 14; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.5e-119;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGTGTAGCTGTAGTGGCTGATTTATATGCTGATTTAATGGGTGATTTGCTTCT 60
DB 1 CATGGTGTAGCTGTAGTGGCTGATTTATATGCTGATTTAATGGGTGATTTGCTTCT 60
QY 61 TCTTTATACCTTTTATTTATCCCAATTTTCTTAAGCAAAATATTTCTTGCTAATCAAT 120
DB 61 TCTTTATACCTTTTATTTATTCCTAAATTTTCTTAAGCAAAATATTTCTTGCTAATCAAT 120
QY 121 AAATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTAT 180
DB 121 AAATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTAT 180
QY 181 CGGGTATATTTTGGAGTTGTAAATATCTACRTGTCTCTTCTTAAGTCCCACTCTCTGT 240
DB 181 CGGGTATATTTTGGAGTTGTAAATATCTACRTGTCTCTTCTTAAGTCCCACTCTCTGT 240
QY 241 TTTCTTTGAGCAGGAAGAAAGCAAGTCACTCTCTTCCATGACAAACCCATGACC 300
DB 241 TTTCTTTGAGCAGGAAGAAAGCAAGTCACTCTCTTCCATGACAAACCCATGACC 300

QY 301 GGCTGCTTGCCTGCTCCATCACAGGAAGGTGAGCAGAGCCCTTTTCTTCCAGGGTTGT 360
DB 301 GGCTGCTTGCCTGCTCCATCACAGGAAGGTGAGCAGAGCCCTTTTCTTCCAGGGTTGT 360
QY 361 GGGACACCTTGAACATATCTGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCA 420
DB 361 GGGACACCTTGAACATATCTGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCA 420
QY 421 AGCTGTTCCGGGACCTTCAGAAATAACCAACACCTGTATTACTGCTTTTGACATCCT 480
DB 421 AGCTGTTCCGGGACCTTCAGAAATAACCAACACCTGTATTACTGCTTTTGACATCCT 480
QY 481 CTCTGGTCCAGGCCCAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 529
DB 481 CTCTGGTCCAGGCCCAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 529

RESULT 3

US-10-027-632-54345
; Sequence 54345, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54345
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-54345

Query Match 99.7%; Score 527.6; DB 13; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-119;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATGGTGATGCTGATGCTGCTGCTGATTTATATGCTGATTTATGCTGATTTTGTCTTCTT 61
DB 1 ATGGTGATGCTGATGCTGCTGCTGATTTATATGCTGATTTATGCTGATTTTGTCTTCTT 60
QY 62 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAAGCAAAATATTTCTTTGCTTAATCAATA 121
DB 61 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAAGCAAAATATTTCTTTGCTTAATCAATA 120
QY 122 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAGAAAGAAAGTTAGCCCTATC 181
DB 121 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAGAAAGAAAGTTAGCCCTATC 180
QY 182 GGGTATATTTTGGAAAGTTGTAAATACTACRTGTTCTCTTCTAAGTCCCACTCTCTCTGT 241
DB 181 GGGTATATTTTGGAAAGTTGTAAATACTACRTGTTCTCTTCTAAGTCCCACTCTCTCTGT 240
QY 242 TTCTTTGAGCGAAGAAAGACGATCACCTCTCTTCTCCATGACAAACCCATGACCG 301
|||||

DB 241 TTCTTTGAGCGAAGAAAGACGATCACCTCTCTTCTCCATGACAAACCCATGACCG 300
QY 302 GCTGCTTGCCTGCTCCATCACAGGAAGGTGAGCAGAGCCCTTTTCTTCCAGGGTTGTG 361
DB 301 GCTGCTTGCCTGCTCCATCACAGGAAGGTGAGCAGAGCCCTTTTCTTCCAGGGTTGTG 360
QY 362 GGCACCTTGAACATATCTGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCAA 421
DB 361 GGCACCTTGAACATATCTGCTGCTAGCTAAGAACTGCCCCAAAGTGAGCACACTCAA 420
QY 422 GCTGTTCCGGGACCTTCAGAAATAACCAACACCTGTATTACTGCTTTTGACATCCTC 481
DB 421 GCTGTTCCGGGACCTTCAGAAATAACCAACACCTGTATTACTGCTTTTGACATCCTC 480
QY 482 TCTGGTCCAGGCCCAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 529
DB 481 TCTGGTCCAGGCCCAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 528

RESULT 4

US-10-027-632-300106
; Sequence 300106, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR APPLICATION NUMBER: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 300106
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-300106

Query Match 99.7%; Score 527.6; DB 13; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-119;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATGGTGATGCTGATGCTGCTGCTGATTTATATGCTGATTTATGCTGATTTTGTCTTCTT 61
DB 1 ATGGTGATGCTGATGCTGCTGCTGATTTATATGCTGATTTATGCTGATTTTGTCTTCTT 60
QY 62 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAAGCAAAATATTTCTTTGCTTAATCAATA 121
DB 61 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAAGCAAAATATTTCTTTGCTTAATCAATA 120
QY 122 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAGAAAGAAAGTTAGCCCTATC 181
DB 121 AATTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAGAAAGAAAGTTAGCCCTATC 180
QY 182 GGGTATATTTTGGAAAGTTGTAAATACTACRTGTTCTCTTCTAAGTCCCACTCTCTCTGT 241
DB 181 GGGTATATTTTGGAAAGTTGTAAATACTACRTGTTCTCTTCTAAGTCCCACTCTCTCTGT 240
QY 242 TTCTTTGAGCGAAGAAAGACGATCACCTCTCTTCTCCATGACAAACCCATGACCG 301
|||||

Db 241 TTCTTTGAGCAGGAAAGAGAGCAGTCAACCTCTCTTCCATGACAAACCCATGACCG 300
QY 302 GCTGCTTGCTGCTCCATCACAGGAGGTGACGAGGCCCTTTCTTCCAGGGTTGTG 361
Db 301 GCTGCTTGCTGCTCCATCACAGGAGGTGACGAGGCCCTTTCTTCCAGGGTTGTG 360
QY 362 GGCACCTTGAACCTATCCTGGCTTGCTAGCTAAGAACTGCCCAAGTGACGACACTCAA 421
Db 361 GGCACCTTGAACCTATCCTGGCTTGCTAGCTAAGAACTGCCCAAGTGACGACACTCAA 420
QY 422 GCTGTTTCGGGACTTCCAGAAATAAACCAACAGCTGTATTACTGCTTTTGACATCCTC 481
Db 421 GCTGTTTCGGGACTTCCAGAAATAAACCAACAGCTGTATTACTGCTTTTGACATCCTC 480
QY 482 TCTGCTCAGGGCCAAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 529
Db 481 TCTGCTCAGGGCCAAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 528

RESULT 5

US-10-027-632-54345
; Sequence 54345, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54345
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-54345

Query Match 99.7%; Score 527.6; DB 14; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-119;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATGGTGATGCTGCTAGTCGCTGATTTATATGCTGATTTATGGTGATTTTGGCTTCCTT 61
Db 1 ATGGTGATGCTGCTAGTCGCTGATTTATATGCTGATTTATGGTGATTTTGGCTTCCTT 60
QY 62 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTGTCTAATCAATA 121
Db 61 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTGTCTAATCAATA 120
QY 122 AATTATCAAAAGAAAAAAGCTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATC 181
Db 121 AATTATCAAAAGAAAAAAGCTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATC 180
QY 182 GGGTATATTTTGGAGTTGTAAAATACTACRTGTTCTCTTCTTAAGTCCCACTCCTCTGTT 241
Db 181 GGGTATATTTTGGAGTTGTAAAATACTACRTGTTCTCTTCTTAAGTCCCACTCCTCTGTT 240
QY 242 TTCTTTGAGCAGGAAAGAGAGCAGTCAACCTCTCTTCCATGACAAACCCATGACCG 301

Db 241 TTCTTTGAGCAGGAAAGAGAGCAGTCAACCTCTCTTCCATGACAAACCCATGACCG 300
QY 302 GCTGCTTGCTGCTCCATCACAGGAGGTGACGAGGCCCTTTCTTCCAGGGTTGTG 361
Db 301 GCTGCTTGCTGCTCCATCACAGGAGGTGACGAGGCCCTTTCTTCCAGGGTTGTG 360
QY 362 GGCACCTTGAACCTATCCTGGCTTGCTAGCTAAGAACTGCCCAAGTGACGACACTCAA 421
Db 361 GGCACCTTGAACCTATCCTGGCTTGCTAGCTAAGAACTGCCCAAGTGACGACACTCAA 420
QY 422 GCTGTTTCGGGACTTCCAGAAATAAACCAACAGCTGTATTACTGCTTTTGACATCCTC 481
Db 421 GCTGTTTCGGGACTTCCAGAAATAAACCAACAGCTGTATTACTGCTTTTGACATCCTC 480
QY 482 TCTGCTCAGGGCCAAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 529
Db 481 TCTGCTCAGGGCCAAAGTCACTGCTCCCTGTCTTGGAGAGGGCTTCTGAG 528

RESULT 6

US-10-027-632-300106
; Sequence 300106, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 300106
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-300106

Query Match 99.7%; Score 527.6; DB 14; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-119;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ATGGTGATGCTGCTAGTCGCTGATTTATATGCTGATTTATGGTGATTTTGGCTTCCTT 61
Db 1 ATGGTGATGCTGCTAGTCGCTGATTTATATGCTGATTTATGGTGATTTTGGCTTCCTT 60
QY 62 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTGTCTAATCAATA 121
Db 61 CTTTATACCTTTTATTTATTTCCAAATTTTCTTAAGCAAAATATTTCTTGTCTAATCAATA 120
QY 122 AATTATCAAAAGAAAAAAGCTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATC 181
Db 121 AATTATCAAAAGAAAAAAGCTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATC 180
QY 182 GGGTATATTTTGGAGTTGTAAAATACTACRTGTTCTCTTCTTAAGTCCCACTCCTCTGTT 241
Db 181 GGGTATATTTTGGAGTTGTAAAATACTACRTGTTCTCTTCTTAAGTCCCACTCCTCTGTT 240
QY 242 TTCTTTGAGCAGGAAAGAGAGCAGTCAACCTCTCTTCCATGACAAACCCATGACCG 301

Db 241 TTTCTTTGAGCAGGAAGAGAAACGAGTCACCCCTCTCTTCCATGACAAACCCATGACGC 300
QY 302 GCTGCTTGCTCTCCATCACAGGAAGTGTCAGAGAGCCCTTTTCTTCCAGGGTTGTG 361
Db 301 GCTGCTTGCTCTCTCATCACAGGAAGTGTCAGAGAGCCCTTTTCTTCCAGGGTTGTG 360
QY 362 GGCACCTTGAACACTCTCTGGCTTGCTAGCTAAAGAACTGCCCAAGTGACGACACTCAA 421
Db 361 GGCACCTTGAACACTCTCTGGCTTGCTAGCTAAAGAACTGCCCAAGTGACGACACTCAA 420
QY 422 GCTGCTCGGGACTTCCAGAAATAACCAACACAGCTGTATTACTGCTTTGACATCCTC 481
Db 421 GCTGCTCGGGACTTCCAGAAATAACCAACACAGCTGTATTACTGCTTTGACATCCTC 480
QY 482 TCTGCTCAGGGCCAAGTCACTGCTCTGCTCTTGGAGAGGGCTTCTGAG 529
Db 481 TCTGCTCAGGGCCAAGTCACTGCTCTGCTCTTGGAGAGGGCTTCTGAG 528

RESULT 7

US-09-814-353-17876
; Sequence 17876, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17876
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match 9.7%; Score 51.4; DB 13; Length 321;
Best Local Similarity 59.4%; Pred. No. 0.011;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 27 TTTATATGCTGATTATGGGTGATTTCCTCTCTTATCTTTATTTATTTCCCAA 86
Db 106 TTTTCTTTAACTTCTTATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 165
QY 87 TTTTCTTTAAAGCAATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACTGA 146
Db 166 TTTTCTTTTAAATTTTCTTTTAAATAAAAAAAATTTAAATAAAAAAAATTTT 225
QY 147 AAGCAACGCTTGAAGGAAA 169
Db 226 ANAAAAAAGAAAAAAGAAAAA 248

RESULT 8

US-09-814-353-17553
; Sequence 17553, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17553
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-814-353-17553

Query Match 9.6%; Score 51; DB 13; Length 284;
Best Local Similarity 61.8%; Pred. No. 0.013;
Matches 81; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 27 TTTATATGCTGATTATGGGTGATTTCCTCTCTTATCTTTATTTATTTCCCAA 86
Db 95 TTTTCTTTTAACTTCTTATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 154
QY 87 TTTTCTTTAAAGCAATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACTGA 146
Db 155 TTTTCTTTTAAATAAAAAATTTTCTTTTAAATAAAAAAAATTTTAAATAAAAAAA 214
QY 147 AAGCAACGCTT 157
Db 215 AAGAAAGAAAT 225

RESULT 9

US-09-960-352-11218/c
; Sequence 11218, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan

Query Match 9.7%; Score 51.4; DB 13; Length 321;
Best Local Similarity 59.4%; Pred. No. 0.011;
Matches 85; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 27 TTTATATGCTGATTATGGGTGATTTCCTCTCTTATCTTTATTTATTTCCCAA 86
Db 106 TTTTCTTTAACTTCTTATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 165
QY 87 TTTTCTTTAAAGCAATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACTGA 146
Db 166 TTTTCTTTTAAATTTTCTTTTAAATAAAAAAAATTTAAATAAAAAAAATTTT 225
QY 147 AAGCAACGCTTGAAGGAAA 169
Db 226 ANAAAAAAGAAAAAAGAAAAA 248

OTHER INFORMATION: Clone ID: 48-LIB3058-026-Q1-K1-D12

US-09-960-352-11218

```
Query Match          9.4%; Score 49.6; DB 10; Length 424;
Best Local Similarity 59.0%; Pred. No. 0.034;
Matches 85; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 26 ATTATATGCTGATTTATGCGGATTTTGGCTCTCTTTATATCTTTTATTTATTTCCCAA 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 AATTTAAATTTATTTATTTTAAATTTTATTTATTTTATTTTATTTTATTTTATTTT 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 86 ATTTTCTTAAGCAATATTTCTTCTATCAATATATCAATATCAATATCAATATCAATATG 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 TTTTTTTTTTTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 146 AAAGCAACGCTTCAAAAAAGGAAA 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 AAAAAAACCACAAAAAAGAAAAA 49
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 10

```
US-09-814-353-16036/c
; Sequence 16036, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-25
; PRIOR FILING DATE: 2000-06-15
; PRIOR FILING DATE: 2000-06-15
; PRIOR FILING DATE: 2000-07-07
; PRIOR FILING DATE: 2000-07-25
; PRIOR FILING DATE: 2000-07-25
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16036
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-16036
```

```
Query Match          9.1%; Score 48.2; DB 13; Length 316;
Best Local Similarity 63.2%; Pred. No. 0.065;
Matches 74; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 53 TCGTTCCTTTATATCTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTCC 112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 TGAGTCCCTTTTTCCTTTGTTTTTCTCTATTTAGTTTATTTAAATTTTATTTTATTTT 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 113 TAATCAATAAATTTATCAAAAGAAAAAAGCAACGCTTGAAAAAGGAAA 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

```
US-10-311-455-2148/c
; Sequence 2148, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: Olek, Alexander
; APPLICANT: Piepenbrock, Christian
```

```
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2148
; LENGTH: 113515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2148
```

```
Query Match          9.0%; Score 47.4; DB 13; Length 113515;
Best Local Similarity 54.4%; Pred. No. 2;
Matches 93; Conservative 1; Mismatches 77; Indels 0; Gaps 0;

QY 55 CTTCCTCTTTATATCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTT 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41662 CTTCCTCTTTATATCAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTT 41603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 115 ATCAATAAATATCAAAAGAAAAAAGCAACGCTTGAAAAAGGAAAGTAG 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41602 ATTTATATATTTATCTATCAATAAAATTTCTAAATTTAAATTTAAATTTAAATTTCA 41543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 175 CCTATCGGTATATTTTGGAGTTGTAATACTACTGTCTCTCTAA 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41542 TTCCACCATTATTTATTTTAAATATCAATTTATCAAAATTTTAACTTATAA 41492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 12

```
US-09-814-353-4872
; Sequence 4872, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4872
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 98, 119, 145, 154, 170, 172, 193, 207, 212, 213, 214,
```


; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18006
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-18006

Query Match 8.8%; Score 46.6; DB 13; Length 383;
Best Local Similarity 56.1%; Pred. No. 0.18; 69; Indels 0; Gaps 0;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 13 GTAGTCGTGGCTGATTATATGCTGATTTATGGGTGATTTGCTTCCTTCTTTACTTT 72
DB 61 GTGTCGCGCGCGAGTACTTTTTTTTTTTTTTTTTTTTTTTTAAAAATTTTTTTTTTT 120
QY 73 TATTTATTCCTCCAAATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATATCAAAA 132
DB 121 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAAATAAAAAA 180
QY 133 GAAAAAAACTCAAGCAACGCTTGAAAAAGGAA 169
DB 181 AAAAAAATAAAAAAATAAAAAAACAACAAAA 217

RESULT 16
US-10-131-827-8192/c
; Sequence 8192, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; FILE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8192
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-8192

Query Match 8.8%; Score 46.4; DB 12; Length 480;
Best Local Similarity 61.7%; Pred. No. 0.22; 46; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 50 TTTTGTCTCTCTTTATCTTTTATTTATTCCTCAATTTTCTTAAGCAATATTTCTT 109
DB 142 TTTTACTTTTTTTTCATATTTTATTAATACCAAAATAAATAAATAAATAATTTTA 83
QY 110 TGTATCAATAAATATCAAGAAAAAATACTGAAAGCAACGCTTGAAAAAGGAA 169

DB 82 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCTTGTAAAAA 23
RESULT 17
US-09-960-352-8414/c
; Sequence 8414, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Weeley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8414
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 36-LIB3058-048-Q1-K1-A8
US-09-960-352-8414

Query Match 8.7%; Score 46.2; DB 10; Length 312;
Best Local Similarity 57.1%; Pred. No. 0.2; 63; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 23 CTGATTATATGCTGATTTATGGGTGATTTGCTTCCTTCTTTATCTTTTATTTC 82
DB 192 CTTTTTTTTTTGTT 133
QY 83 CAAATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATATCAAGAAAAA 142
DB 132 TTTTTTTTTTTTTTTTTTTTATTAATAAAAAAATAAAAAAATAAAAAA 73
QY 143 CTGAAAGCAACGCTTGAAAAAGGAA 169
DB 72 AAAAAAAAAAACAACAAAAA 46

RESULT 18
US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 8.7%; Score 46; DB 13; Length 3673778;
Best Local Similarity 57.7%; Pred. No. 25; 60; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 23 CTGATTATATGCTGATTTATGGGTGATTTGCTTCCTTCTTTATCTTTTATTTC 82
DB 88260 CCGACTCTTACTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATATTTATCATCT 88201
QY 83 CAAATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATATCAAGAAAAA 142

Db 88200 TAAATATTCTCAAAAAAATAATACAAATCATATAATAATAATAAAAAAATC 88141
QY 143 CTGAAGCAACGCTTGAAAAA 164
Db 88140 AACAAATAACACGTAAACAA 88119

RESULT 19
US-09-814-353-4849
; Sequence 4849, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4849
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 57, 58, 59, 60, 100, 102, 128, 129, 130, 131, 132, 133, 137,
; LOCATION: 138, 140, 141, 142, 143, 150, 174, 176, 177, 184, 185, 190,
; LOCATION: 198, 199, 208, 209, 210, 211, 212, 213, 219, 222, 232, 234,
; LOCATION: 249, 252, 253, 254, 263, 264, 265, 268, 269, 276
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 284, 286, 294, 295, 296, 297, 298, 299, 300, 301, 302, 311,
; LOCATION: 313, 314, 325, 326, 335, 340, 344, 345, 346, 349, 350, 351,
; LOCATION: 359, 367, 379, 382, 386, 387, 389, 394, 395, 403, 406, 408,
; LOCATION: 409, 413, 414, 424, 426, 428, 429, 432, 433, 437, 438
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 439, 440, 441, 442, 444, 447, 451
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4849

Query Match 8.7%; Score 45.8; DB 13; Length 463;
Best Local Similarity 44.9%; Pred. No. 0.3;
Matches 89; Conservative 1; Mismatches 108; Indels 0; Gaps 0;
QY 67 TACTTTTATTTCCCAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATTA 126
Db 26 TACTTT 85
QY 127 TCAAAAGAAAAAACTGAAGCAACGCTTGAAAAAGAAAGTTAGCCCTATCGGTA 186
Db 86 TAAAAAATAAANNCNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 145
QY 187 TATTTTGAAGTTGTAAATACTACRTGTCTCTTCTAAGTCCCACTCCTCTGTTCTT 246

Db 146 TTTTNGGGGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 205
QY 247 TGACGAGGAAAGAGAAAG 264
Db 206 TTNNNNNAAAAAANAANG 223
RESULT 20
US-09-814-353-11146
; Sequence 11146, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11146
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 57, 58, 59, 60, 100, 102, 128, 129, 130, 131, 132, 133, 137,
; LOCATION: 138, 140, 141, 142, 143, 150, 174, 176, 177, 184, 185, 190,
; LOCATION: 198, 199, 208, 209, 210, 211, 212, 213, 219, 222, 232, 234,
; LOCATION: 249, 252, 253, 254, 263, 264, 265, 268, 269, 276
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 284, 286, 294, 295, 296, 297, 298, 299, 300, 301, 302, 311,
; LOCATION: 313, 314, 325, 326, 335, 340, 344, 345, 346, 349, 350, 351,
; LOCATION: 359, 367, 379, 382, 386, 387, 389, 394, 395, 403, 406, 408,
; LOCATION: 409, 413, 414, 424, 426, 428, 429, 432, 433, 437, 438
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 439, 440, 441, 442, 444, 447, 451
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-11146

Query Match 8.7%; Score 45.8; DB 13; Length 463;
Best Local Similarity 44.9%; Pred. No. 0.3;
Matches 89; Conservative 1; Mismatches 108; Indels 0; Gaps 0;
QY 67 TACTTTTATTTCCCAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATTA 126
Db 26 TACTTT 85
QY 127 TCAAAAGAAAAAACTGAAGCAACGCTTGAAAAAGAAAGTTAGCCCTATCGGTA 186
Db 86 TAAAAAATAAANNCNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 145
QY 187 TATTTTGAAGTTGTAAATACTACRTGTCTCTTCTAAGTCCCACTCCTCTGTTCTT 246

Query Match	8.5%	Score 45;	DB 13;	Length 359;
Best Local Similarity	53.0%;	Pred. No. 0.42;		
Matches	96;	Conservative	0;	Mismatches 85; Indels 0; Gaps 0
Qy	12	TCTAGTCGGCGTATTATATGCTGATTTATATGGGTCATTTTCCTTCCTTTTATACCT	71	


```
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11024
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 56, 57, 59, 137, 147, 148, 156, 162, 167, 169, 173, 189,
; LOCATION: 200, 202, 214, 223, 225, 227, 231, 239, 240, 248, 257, 259,
; LOCATION: 260, 266, 270, 273, 274, 275, 277, 278, 280, 283, 287, 295,
; LOCATION: 306, 311, 313, 317, 340, 344, 351, 353, 361, 367, 370
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 371, 378, 379, 381, 400, 403, 406, 407, 416, 421, 428, 431,
; LOCATION: 434, 475, 476, 479, 480, 482, 483, 491, 515, 518, 520, 522,
; LOCATION: 528, 529, 533, 535, 536, 558, 561, 578, 581, 586, 588, 589,
; LOCATION: 593, 597, 603, 615, 618, 627, 639, 643, 645, 653, 658
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 686, 688, 689, 690, 692, 694, 695, 696, 697, 698, 699, 700,
; LOCATION: 701, 703, 704, 706, 711, 713, 714
; OTHER INFORMATION: n = A,T,C or G
; US-09-814-353-11024

Query Match 8.5%; Score 44.8; DB 13; Length 723;
Best Local Similarity 57.9%; Pred. No. 0.67;
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 49 ATTTGCTTCCTCTTATATCTTTTATTTATTCCTCAAAATTTTCTTAAGCAAAATTTCT 108
Db 191 AANTTTTTTTTTTTTTTTTCCNCNTTTTTTTTAAATAATTTTNNAAAAAATNTTTTT 132

Qy 109 TTGCTAATCAATAAATATCAAAAGAAAAAACTGAAGCAACGCTTGAAGAAAGGAA 168
Db 131 TTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAA 72

Qy 169 A 169
Db 71 A 71

RESULT 27
US-09-814-353-17588/c
; Sequence 17588, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17588
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-814-353-17588

Query Match 8.4%; Score 44.4; DB 13; Length 228;
Best Local Similarity 59.5%; Pred. No. 0.47;
Matches 75; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 45 GGTGATTTTGCTTCCTCTTATATCTTTTATTTATTCCTCAAAATTTTCTTAAGCAAAAT 104
Db 227 GGTGTTTTTTTATCTCTCTTTTTTTTTTTTATCTTTTATCTTTTATCTTTTATCTTTT 168

Qy 105 TTCTTTGCTAATCAATAAATATCAAAAGAAAAAACTGAAGCAACGCTTGAAGAAAA 164
Db 167 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAAT 108

Qy 165 GGAAAG 170
Db 107 AAAAAAG 102

RESULT 28
US-09-834-975-451
; Sequence 451, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 451
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)...(425)
; OTHER INFORMATION: n = A,T,C or G
; US-09-834-975-451

Query Match 8.4%; Score 44.4; DB 10; Length 425;
Best Local Similarity 54.9%; Pred. No. 0.64;
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 27 TTTATATGCTGATTTTAAAGGATTTTGTCTTCTTCTTATATCTTTTATCTTTTATCTTTT 86
```


	Query Match	8.4%	Score 44.4;	DB 13;	Length 6078;	
	Best Local Similarity	54.4%;	Pred. No. 2.5;	Mismatches 1;	Gaps 0;	
	Matches 87;	Conservative	72;	Indels 0;	Gaps 0;	
Qy	65	TATACITTTTATTTATTC	CCAAAATTTTCTT	TAAGCAAAATATTTCTTTGCTCAATCAATAAAT	124	
Db	4494	TAAAAATATACCAAAATCTT	TAAATATATATAAAAAA	AAAAAACCTCACACATAAATAAAA	4435	
Qy	125	TATCAAAAGAAAAA	AAAAAAGCTGAAAGCAAGCGTCTG	AAAAAAGGAAGTAGCCCTCATCGGG	184	
Db	4434	AAAAA	AAAAAACAATATTA	AAAAACAACCTTCATATAATATATCTT	AAATACCTTTT	4375
Qy	185	TATATTTTGAAGTCTG	TAAATACTACRTGTCTCTCTCA	224		
Db	4374	TATTTTATATTTTATATTTT	TATTTTCCAAATTTCTACAA	4335		

```

RESULT 34
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2

```

```

; NUMBER OF SEQ ID NOS: 2
;
; SEQ ID NO 1
;
; LENGTH: 3673778
;
; TYPE: DNA
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;
; FEATURE:
;
; NAME/KEY: unsure
;
; LOCATION: (3294164)
;
;
US-10-312-841-1

Query Match      8.4%; Score 44.2; DB 13; Length 3673778;
Best Local Similarity 51.8%; Pred. No. 68;
Matches 100; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy      17  TCGTGGCTGATTATATGCTGATTATGCGTGATTTGCTTCCTCTTATACCTTTATT 76
Db      686838  TGGTTTTTGGTGTGCTTTTGGATTATCGTTATTTTAAAGTTTTTTTTTTAGTATTGATA 686838

```

[illegible]

```
QY 197 GTGTAAATACT 209
|||
Db 687018 ATTTAAATAAT 687030

RESULT 35
US-09-814-353-17383
; Sequence 17383, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-005B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17383
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-17383

Query Match 8.3%; Score 44; DB 13; Length 281;
Best Local Similarity 57.1%; Pred. No. 0.65;
Matches 80; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTATGCGTATTTGCTTCTCTTTATCTTTTATCTTTATCTTCCAAA 86
|||
Db 141 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 200
|||

QY 87 TTTTCTTAAAGCAATATTTCTTTGCTAATCAATAAATATCAAAAGAAAAAACTGA 146
|||
Db 201 TTTTAAAAAATAATTTTAAATTAATAAATAATTTCAATTAATAAATAAATAAATCTTT 260
|||

QY 147 AAGCAGCTTGAAAAAGG 166
|||
Db 261 ATTGAAAAAATAAAGG 280
|||

RESULT 36
US-10-198-846-8434
; Sequence 8434, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8434
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2, 14, 27, 61, 150, 153, 156, 159, 160, 218, 222, 226, 227,
; LOCATION: 229, 230, 232, 234, 236, 288, 289, 290, 291, 292, 316, 319,
; LOCATION: 323, 324, 325, 326, 328, 330, 331, 332, 333, 334, 340, 345,
; LOCATION: 346, 349, 350, 351, 352, 368, 369, 370, 371, 372, 373
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 374, 375, 376, 377, 379, 381, 393, 395, 396, 397, 399, 400,
; LOCATION: 401, 402, 403, 404, 406, 408, 411, 422, 423, 424, 425, 426,
; LOCATION: 427, 435, 436, 437, 444, 445, 448, 449, 469, 472, 473,
; LOCATION: 478, 479, 480, 481, 483, 497, 502, 505, 506, 514, 517
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 518, 519, 520, 521, 522, 524, 526, 527, 528, 533, 534, 535,
; LOCATION: 536, 537, 538, 541, 559, 561, 562, 563, 564, 567, 568, 569,
; LOCATION: 570, 575, 589, 590, 592, 594, 598, 599
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-8434

Query Match 8.3%; Score 44; DB 15; Length 600;
Best Local Similarity 48.7%; Pred. No. 0.96;
Matches 95; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 13 GTAGTCGTGGCTGATTATATGCTGCTGATTATGCGTGATTTCCTTCCTTCTTTACTTT 72
|||
Db 60 GNACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 119
|||

QY 73 TATTATTCCCAAAATTTTCTTAAGCAATATTTCTTGTCTAATCAATAAATTTATCAAAA 132
|||
Db 120 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 179
|||

QY 133 GAAAAAAACGTGAAGCAACGCTTGAAAAAGAAAAAGTACCCCTATCGGTATATTTT 192
|||
Db 180 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 239
|||

QY 193 CGAAGTGTGAATA 207
|||
Db 240 GGGGAAAAAATAAATA 254
|||

RESULT 37
US-09-960-352-12183
; Sequence 12183, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12183
; LENGTH: 236
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 52-LIB3058-016-Q1-K1-E12
US-09-960-352-12183

Query Match 8.3%; Score 43.8; DB 10; Length 236;
```


Best Local Similarity 56.6%; Pred. No. 0.67;
Matches 81; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTATGGGTGATTTTGGCTCTCTTATACCTTTATTTATTTATCCAAA 86
DB 47 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 106

QY 87 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 146
DB 107 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 166

QY 147 AAGCAACGCTTGAAAAAGAAA 169
DB 167 AATGTTAGTGTGAAAAATGACAAA 189

RESULT 38

US-09-814-353-17383/c
; Sequence 17383, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-814-353-17383
Query Match 8.3%; Score 43.8; DB 13; Length 281;
Best Local Similarity 56.6%; Pred. No. 0.73; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 27 TTATATGCTGATTATGGGTGATTTTGGCTCTCTTATACCTTTATTTATTTATCCAAA 86
DB 275 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 216
QY 87 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 146
DB 215 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 156
QY 147 AAGCAACGCTTGAAAAAGAAA 169
DB 155 AAAAAAAGAAAAAAGAAAAA 133

RESULT 39

US-09-814-353-5380
; Sequence 5380, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John

APPLICANT: Thompson, Pamela
APPLICANT: Lillie, James
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5380
LENGTH: 610
TYPE: DNA
ORGANISM: Homo sapiens

NAME/KEY: misc feature
LOCATION: 50, 99, 103, 104, 105, 106, 107, 108, 112, 114, 120, 121,
LOCATION: 122, 127, 128, 131, 132, 133, 134, 150, 151, 152, 153, 154,
LOCATION: 155, 156, 157, 158, 159, 160, 161, 162, 163, 183, 184, 186,
LOCATION: 188, 189, 190, 192, 193, 195, 196, 197, 209, 221, 223
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc feature
LOCATION: 224, 225, 238, 245, 248, 255, 257, 258, 260, 261, 269, 276,
LOCATION: 277, 278, 279, 280, 289, 292, 293, 302, 308, 322, 330, 335,
LOCATION: 343, 346, 348, 349, 350, 353, 365, 366, 367, 371, 377, 384,
LOCATION: 385, 398, 408, 422, 423, 440, 452, 453, 454, 455, 456
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc feature
LOCATION: 457, 458, 459, 461, 473, 474, 475, 476, 480, 487, 494,
LOCATION: 497, 512, 521, 527, 540, 543, 544, 547, 549, 550, 554, 568,
LOCATION: 569, 584, 586, 587, 589, 591, 595, 609
OTHER INFORMATION: n = A,T,C or G

US-09-814-353-5380
Query Match 8.3%; Score 43.8; DB 13; Length 610;
Best Local Similarity 36.8%; Pred. No. 1.1;
Matches 120; Conservative 1; Mismatches 205; Indels 0; Gaps 0;
QY 27 TTATATGCTGATTATGGGTGATTTTGGCTCTCTTATACCTTTATTTATTTATCCAAA 86
DB 34 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 93
QY 87 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 146
DB 94 TTTTCTTAAAGCAAAATTTCTTGGCTAAATTAATCAAAAGAAAAAACTGA 153

QY 147 AAGCAACGCTTGAAAAAGAAAAGTTAGCCCTATCGGGTATATTTTGGAGTTGTAAAT 206
DB 154 NNN 213
QY 207 ACTACTGTCTCTTCTTAAGTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 266
DB 214 AAAAAAANGNNNTTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 273
QY 267 GTCACCTCTCTTCCATGACCAACCCATGACCGGCTGCTGCTGCTGCTGCTGCTGCT 326
DB 274 TNNNNCCCTTTTNGGNN 333

QY 327 AGGTGCAGCAGAGCCCTTTTCTTCCC 352

US-10-311-455-1449

Query Match	8.2%	Score 43.6	DB 13	Length 6590
Best local Similarity	57.2%	Pred. No. 4		
Matches 79	Conservative 0	Mismatches 59	Indels 0	Gaps 0
Qy	27	TTTTATGCTGATTTATGGGTGATTTCCTCTCTTTATACITTTTATTATTTATTTCCCAA	86	
Db	1088	TTTTCTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTATTAAATCATCTCTTAA	1029	
Qy	87	TTTTTCTTAAGCAAAATATTTCTTTCGCTAAATCAATAAATTCAAAAAGAAAAAAACTGA	146	
Db	1028	TATTTCTCGAAAAAATAATAATAAATCATAAATATAATAAATAAATAAATAAATCAACA	969	
Qy	147	AAGCAACGCTTGAAGAAA	164	
Db	968	AATAAACACGTAAACAAA	951	

RESULT 43

US-09-960-352-11218
 ; Sequence 11218, Application US/09960352
 ; Patent No. US20020137139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathialagan, Nagappan
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 ; FILE REFERENCE: 16511.006/37-21(10298)C
 ; CURRENT APPLICATION NUMBER: US/09/960,352
 ; CURRENT FILING DATE: 2001-09-24
 ; NUMBER OF SEQ ID NOS: 15112
 ; SEQ ID NO 11218
 ; LENGTH: 424
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; OTHER INFORMATION: Clone ID: 48-LIB3058-026-Q1-K1-D12
 US-09-960-352-11218

Query Match	8.2%	Score 43.4	DB 10	Length 424
Best Local Similarity	49.3%	Pred. No. 1.1		
Matches 110	Conservative 1	Mismatches 112	Indels 0	Gaps 0
Qy	4	GGTGATCGCTGAGTCGGCTGATTATATGCTGATTTATGCGGTGATTTGCTCTCTCTCTCT	63	
Db	10	GGTCGCCCTTT	69	
Qy	64	TTATACATTTTATTTATTTCCCAAAATTTTTCTTAAGCAAAATATTTCTTTGCTAATCAATAAA	123	
Db	70	TTTAAATATAAAAAATAAAAAAAA	129	
Qy	124	TTATCAAAAGAAAAAAAACCTGGAAGCAACGCTTGAAAAAGGAAGTAGCCCTATCGG	183	
Db	130	AAATAAAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAATATAA	189	
Qy	184	GTATATTTTGGAGTGTGAAATACTACRTGTTCTCTCTTAAG	226	
Db	190	ATTTTAAATATTTCTTTTAAATATAAAAAATAAAATATTTATTAAG	232	

RESULT 44

US-09-960-352-9335
; Sequence 9335, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

```

; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 9335
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 40-LIB3058-035-Q1-K1-B8
US-09-960-352-9335

```

Query Match	8.2%	Score 43.2;	DB 10;	Length 380;
Best Local Similarity	56.2%;	Pred. No. 1.2;		
Matches	81;	Conservative 0;	Mismatches 63;	Indels 0; Gaps 0;
Qy	27	TTTTATGCTGATTTATGGTGATTTTGGCTCTCTTTATACGTTTTATTTATTTCCAAA	86	
Db	49	TTTTTGGTTGTTT	108	
Qy	87	TTTTTCTTAAGCAAAATATTTCTTTGGCTAATCAATAAATATCAAAAGAAAAAACTGA	146	
Db	109	TTTTTTTTTATTTATAATAATATTTTCAATTTAAAAAATATAATGTTTAAAAATTTTTT	168	
Qy	147	AAGCAACGCTTGAAAAAAGGAAAG	170	
Db	169	AAAAAAAATTTAAAAAATAAAAAAAG	192	

RESULT 45

```

US-09-814-353-17565/c
; Sequence 17565, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Little, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: WRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17565
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-17565

```

	Query Match	8.2%	Score 43.2	DB 13	Length 397
	Best Local Similarity	54.4%	Pred. No. 1.2		
	Matches 87	Conservative 0	Mismatches 73	Indels 0	Gaps 0
Qy	15	AGTCGGCGGCGATTTATATGCTGATTATGGGCGATTTTGGCTTCCTCTCTTATACCTTTTA	74		
Db	267	AGTCCTTTGGGATGCTTTTTTAAATGTTTGTTTTTCTCTTTATTTTTTTTTTCTCA	208		
Qy	75	TTTATTTCCCAATTTTTCCTAAGCAAAATATTCCTTGGCTTAATCAATTAATATCAAAAGA	134		
Db	207	TTTTTTTTTTTTTAATCCCTCTTAAAAAATTTTTTTTTTTTAACTTTAAATAATAAAAAGAAAAA	148		

Qy 135 AAAAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTTAG 174
Db 147 AAAAAAAAAACAAAAAAGAAAAAAGAAAAAAGAAAAAAG 108

Search completed: February 14, 2004, 14:16:16
Job time : 357 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 12:05:58 ; Search time 277 Seconds
(without alignments)

5155.244 Million cell updates/sec

Title: US-10-027-632-1

Perfect score: 529

Sequence: 1 catggtgatgtgtagtctg.....tcttgagagggtctctgag 529

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : N Geneseq 15Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50.4	9.5	556	23	Human prostate exp
2	50.4	9.5	556	23	Human prostate exp
3	50.4	9.5	556	23	Human prostate exp
4	50.4	9.5	556	23	Human prostate exp
5	49.6	9.4	424	25	Bovine EST associa
6	49.2	9.3	5852	12	Dictyostelium plas
7	48.6	9.2	464	23	Human prostate exp
8	48.6	9.2	612	22	Human cervical can

9	48	9.1	474	23	ABV56492	Human prostate exp
10	47.8	9.0	300	23	ABV04310	Human prostate exp
11	47.4	9.0	113515	24	ABL34175	Human immune syste
12	47	8.9	309	23	ABV44994	Human prostate exp
13	47	8.9	309	23	ABV44994	Human prostate exp
14	47	8.9	474	23	ABV56492	Human prostate exp
15	47	8.9	612	22	AAH71471	Human cervical can
16	46.6	8.8	312	25	ABX43249	Bovine EST associa
17	46.4	8.8	480	24	ABZ08201	Human leukocyte de
18	46.2	8.7	312	25	ABX43249	Bovine EST associa
19	45.4	8.6	556	23	ABV40063	Human prostate exp
20	45.4	8.6	556	23	ABV40163	Human prostate exp
21	45.4	8.6	556	23	ABV42105	Human prostate exp
22	45.4	8.6	556	23	ABV43601	Human prostate exp
23	44.6	8.4	464	23	ABV03810	Human prostate exp
24	44.6	8.4	10988	24	ABL70200	Chemically treated
25	44.6	8.4	10988	24	AA561158	Human gene regulat
26	44.6	8.4	10988	24	ABK31245	Signal transductio
27	44.4	8.4	425	22	AA560450	Human cancer agent
28	44.4	8.4	494	23	ABV10021	Human prostate exp
29	44.4	8.4	664	21	AAA26336	Human secreted pro
30	44.4	8.4	6078	22	AAA46406	Tumour suppressor
31	44.4	8.4	6078	24	ABL33137	Human immune syste
32	43.8	8.3	198	22	AA507727	Cervical cancer pr
33	43.8	8.3	236	25	ABX47018	Bovine EST associa
34	43.8	8.3	323	23	ABV60996	Human prostate exp
35	43.8	8.3	378	22	AAI90863	Human polynucleoti
36	43.8	8.3	411	22	AAI88595	Human polynucleoti
37	43.8	8.3	9884	24	ABK33925	Human DNA for stag
38	43.8	8.3	56153	22	AA546793	Tumour suppressor
39	43.6	8.2	626	23	ABV60941	Human prostate exp
40	43.6	8.2	6590	24	ABL33476	Human immune syste
41	43.4	8.2	300	22	AAH70049	Human cervical can
42	43.4	8.2	424	25	ABV46053	Bovine EST associa
43	43.4	8.2	431	23	ABV04399	Human prostate exp
44	43.2	8.2	380	25	ABX44170	Bovine EST associa
45	43.2	8.2	430	24	ABV83644	Human breast speci
46	43.2	8.2	12426	22	AA546481	Tumour suppressor
47	43	8.1	255	22	AAH82206	Rat differential t
48	43	8.1	180557	24	ABN85750	Human BAC clone Rp
49	43	8.1	368004	24	ABL57909	Human transporter
50	42.8	8.1	404	23	ABV56394	Human prostate exp
51	42.6	8.1	522	22	AA534084	Human cDNA encodin
52	42.6	8.1	621	22	AAH71551	Human cervical can
53	42.6	8.1	768	22	AAAL08093	Human breast cance
54	42.4	8.0	308	25	ABX42505	Bovine EST associa
55	42.4	8.0	445	23	ABV05431	Human prostate exp
56	42.4	8.0	524	25	ABT22885	Breast cancer mark
57	42.4	8.0	2672	19	AAV39862	Mycoplasma hyopneu
58	42.4	8.0	6731	24	ABL33060	Human immune syste
59	42.4	8.0	277	25	ABX47508	Bovine EST associa
60	42.2	8.0	277	25	ABX47508	Bovine EST associa
61	42.2	8.0	294	22	AAH70082	Human cervical can
62	42.2	8.0	380	25	ABX44170	Bovine EST associa
63	42.2	8.0	397	22	AAI82664	Human polynucleoti
64	42.2	8.0	411	22	AAI88595	Human polynucleoti
65	42.2	8.0	418	25	ABX39680	Bovine EST associa
66	42.2	8.0	453	23	ABV09564	Human prostate exp
67	42.2	8.0	493	22	AAI25179	Human breast cance
68	42.2	8.0	550	23	ABV57072	Human prostate exp
69	42.2	8.0	1066	24	ABN98373	Arabidopsis thalia
70	42.2	8.0	5359	24	ABL33607	Human immune syste
71	42.2	8.0	5359	24	ABK28272	DNA transcription
72	42.2	8.0	6533	24	ABK28383	DNA transcription
73	42	7.9	1342	21	AAA26424	Human secreted pro
74	42	7.9	4163	24	ABL34330	Human immune syste
75	42	7.9	13449	24	ABL33384	Human immune syste
76	41.8	7.9	389	22	AAAL20088	Human breast cance
77	41.8	7.9	396	22	AAAF94842	Human ovarian can
78	41.8	7.9	396	24	ABT03109	Human ovarian carc
79	41.8	7.9	396	24	ABL48792	Ovarian carcinoma
80	41.8	7.9	612	22	AAH71474	Human cervical can
81	41.8	7.9	2001	18	AAAT59700	PTH-like peptide D

c 82 41.8 7.9 5020 25 ABZ09960
 c 83 41.8 7.9 513445 22 AA161373
 c 84 41.6 7.9 198 22 AAS07727
 c 85 41.6 7.9 298 25 ABX35839
 c 86 41.6 7.9 545 22 AAH70126
 c 87 41.6 7.9 566 20 AAY86948
 c 88 41.6 7.9 693 6 AAN50336
 c 89 41.6 7.9 693 11 AAQ06408
 c 90 41.6 7.9 1180 24 AAD33724
 c 91 41.6 7.9 6794 24 AAB70220
 c 92 41.6 7.9 6794 24 AAS61175
 c 93 41.6 7.9 6794 24 ABK31265
 c 94 41.6 7.9 22073 22 AAL04663
 c 95 41.6 7.9 22073 23 ABL97570
 c 96 41.6 7.9 22073 25 ABZ74560
 c 97 41.6 7.9 22073 25 ABZ68092
 c 98 41.4 7.8 418 25 ABX39680
 c 99 41.4 7.8 643 22 AAL16045
 c 100 41.4 7.8 1617 24 ABS59408

ALIGNMENTS

RESULT 1
 ABV40063
 ID ABV40063 standard; cDNA; 556 BP.
 AC ABV40063;
 DT 16-SEP-2002 (first entry)
 DE Human prostate expression marker cDNA 40054.
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 PN WO200160860-A2.
 PD 23-AUG-2001.
 PF 20-FEB-2001; 2001WO-US05171.
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX Claim 1; Page 8100; 11750pp; English.
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (i) is also useful as a pharmacodynanamic or pharmacogenomic marker.
 XX Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;
 SQ
 Query Match 9.5%; Score 50.4; DB 23; Length 556;
 Best Local Similarity 58.8%; Pred. No. 0.02;
 Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 22 GCTGATTTATATGCTGATTATGCGTGGATTGCTTCTTCTTATCTTTATTATTC 81
 DB 1 GCTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 60
 QY 82 CCAATTTTCTTAAGCAATAATTTCTTTGCTAATCAATAATTAATCAAGAAAAAAA 141
 DB 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 120
 QY 142 ACTGAAGCAACGCTTGAAGAAAGGAAA 169
 DB 121 AAAAAAAGAAAAAAGAAAAAAGAAAAA 148

RESULT 2
 ABV40163
 ID ABV40163 standard; cDNA; 556 BP.
 AC ABV40163;
 DT 16-SEP-2002 (first entry)
 DE Human prostate expression marker cDNA 40154.
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 PN WO200160860-A2.
 PD 23-AUG-2001.
 PF 20-FEB-2001; 2001WO-US05171.
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX Claim 1; Page 8115; 11750pp; English.
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;

Query Match 9.5%; Score 50.4; DB 23; Length 556;
Best Local Similarity 58.8%; Pred. No. 0.02;
Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 GCTGATTTATGCTGATTTATGGTGATTTGCTTCCTCTTATACCTTTATTATTTC 81
Db 1 GCTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 60

Qy 82 CCAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 141
Db 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120

Qy 142 ACTGAAGCAACGCTTGAAAAAGAAA 169
Db 121 AAAAAAAGAAAAAAGAAAAAAGAAAAA 148

RESULT 3
ABV42105
ID ABV42105 standard; cDNA; 556 BP.
XX
AC ABV42105;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 42096.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC..
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 8441; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 2 other;

Query Match 9.5%; Score 50.4; DB 23; Length 556;
Best Local Similarity 58.8%; Pred. No. 0.02;
Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 GCTGATTTATGCTGATTTATGGTGATTTGCTTCCTCTTATACCTTTATTATTTC 81
Db 1 GCTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 60

Qy 82 CCAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAAATATCAAAAGAAAAA 141
Db 61 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 120

Qy 142 ACTGAAGCAACGCTTGAAAAAGAAA 169
Db 121 AAAAAAAGAAAAAAGAAAAAAGAAAAA 148

RESULT 4
ABV43601
ID ABV43601 standard; cDNA; 556 BP.
XX
AC ABV43601;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 43592.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
XX
PR 25-MAY-2000; 2000US-207454P.
XX
PR 09-JUN-2000; 2000US-211314P.
XX
PR 18-JUL-2000; 2000US-219007P.
XX
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC..
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 8681-8682; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
CC specification or its complement. (I) is useful for:

XX PA (UTMA-) UNIV MACQUARIE.
 XX PI Slade MB, Chang ACM, Williams KL;
 XX DR WPI; 1991-164194/22.
 XX DR P-PSDB; AAR11988.
 XX PT Polypeptide facilitating extra-chromosomal replication - of
 XX PT recombinant plasmid in Dictyostelium species
 XX PS Claim 15; Fig 1; 90pp; English.
 XX CC The sequence of Ddp2 has been found to contain the putative open
 CC reading frame indicated in the Features Table. The possible ORF is
 CC flanked by regions with similarity to promoter and poly adenylation
 CC signals of known Dictyostelium genes. The RNA and polypeptide
 CC product of the Rep gene have not, however, been detected. It is
 CC believed that the product is produced in low amounts to positively
 CC regulate initiation of plasmid replication. The polypeptide may also
 CC contain regions that act as negative regulators of plasmid copy
 CC number. See also AAQ11711 and AAQ11712.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX CC
 XX SQ Sequence 5852 BP; 2298 A; 651 C; 708 G; 2195 T; 0 other;
 Query Match 9.3%; Score 49.2; DB 12; Length 5852;
 Best Local Similarity 57.0%; Pred. No. 0.076; 68; Indels 0; Gaps 0;
 Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 12 TGTAGTCGTGGCTGATTTATATGCTGATTTATGGGTGATTTTGGCTTCTTATACCTT 71
 DB 5767 TTTAATTATTTGTTATTTGTTATTTTATATATGTTATTTGTTGTTTACTT 5708
 QY 72 TTAATTATCCCAATTTTCTTAAGCAATATTTCTTGCTAATCAATAATATCAAA 131
 DB 5707 ATATTTCTATTTTATTTATTAATAATTAATTAATTAATTAATAATAAAAAA 5648
 QY 132 AGAAAAAACTGAAGCAAGCTTGAAAAAGGAAA 169
 DB 5647 AAAAAAAATTTAATTAATAATAATAATAATAATAATAAAAAA 5610
 RESULT 7
 ABV03810
 ID ABV03810 standard; cDNA; 464 BP.
 XX AC ABV03810;
 XX DT 13-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 3801.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US05171.
 XX PR 17-FEB-2000; 2000US-183319P.
 XX PR 16-MAR-2000; 2000US-189862P.
 XX PR 25-MAY-2000; 2000US-207454P.
 XX PR 09-JUN-2000; 2000US-211314P.
 XX PR 18-JUL-2000; 2000US-219007P.
 XX PR 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PA
 XX

PI Schlegel R, Endege WO, Monahan JE;
 XX DR WPI; 2001-662795/76.
 XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 XX PT prostate cells and correlating with presence of prostate cancer, useful
 XX PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX PS Claim 1; Page 677; 11750pp; English.
 XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX SQ
 Query Match 9.2%; Score 48.6; DB 23; Length 464;
 Best Local Similarity 43.3%; Pred. No. 0.05; 172; Indels 0; Gaps 0;
 Matches 132; Conservative 1; Mismatches 172; Indels 0; Gaps 0;
 QY 27 TTTATATGCTGATTTATGGGTGATTTTGGCTTCTTATACCTTTTATTTATTTCCCAA 86
 DB 8 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 67
 QY 87 TTTTCTTAAGCAATATTTCTTTGCTTAATCAATTAATTAATTAATTAATTAATTAATTA 146
 DB 68 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 127
 QY 147 AAGCAAGCTTGAAAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAGTTGTAAT 206
 DB 128 AAAAAAAATTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 187
 QY 207 ACTACRTGTTCTCTTAAGTCCCACTCTCTGTTTCTTTGAGCAGAGAGAGAGAGCA 266
 DB 188 AAAAAAAATTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 247
 QY 267 GTCACCTCTCTTCCATGACAAACCCATGACCGGCTGCTTGCCTGCTCATCAGGA 326
 DB 248 AACCCCNCCCAANNNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 307
 QY 327 AGGTG 331
 DB 308 AGGGG 312
 RESULT 8
 AAH71471
 ID AAH71471 standard; cDNA; 612 BP.
 XX AC AAH71471;
 XX DT 19-SEP-2001 (first entry)
 XX DE Human cervical cancer marker nucleic acid 2745.
 XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN WO200142467-A2.
 XX

```
PD 14-JUN-2001.
XX
XX 08-DEC-2000; 2000MO-US33312.
XX
XX 08-DEC-1999; 99US-0169681.
XX 21-DEC-1999; 99US-0171350.
XX 14-MAR-2000; 2000US-0189315.
XX 12-MAY-2000; 2000US-0203791.
XX 09-JUN-2000; 2000US-0210600.
XX 21-JUL-2000; 2000US-0220114.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
XX and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 564; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy.
XX
XX SQ Sequence 612 BP; 380 A; 50 C; 2 G; 178 T; 2 other;
XX
XX Query Match 9.2%; Score 48.6; DB 22; Length 612;
XX Best Local Similarity 58.7%; Pred. No. 0.054;
XX Matches 84; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
XX
QY 27 TTATATATGCTGATTTATGGTGATTTTGTCTTCTCTTATATCTTTTATTTATTCCTCAA 86
Db 17 TTTTCTTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 76
QY 87 TTTTCTTAAAGCAATATTTCTTTGCTAATCAATAATATCAAAAGAAAAAACTGCA 146
Db 77 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 136
QY 147 AAGCAAGCTTGAAGAGGAA 169
Db 137 AAAAAAAATAAAAAAA 159
XX
RESULT 9
ABV56492
ID ABV56492 standard; cDNA; 474 BP.
XX
XX ABV56492;
XX
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 56483.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX
```

```
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JB;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 10898; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX SQ Sequence 474 BP; 234 A; 34 C; 107 G; 98 T; 1 other;
XX
XX Query Match 9.1%; Score 48; DB 23; Length 474;
XX Best Local Similarity 59.6%; Pred. No. 0.07;
XX Matches 81; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
XX
QY 34 GCTGATTTATGGTGATTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 93
Db 9 GCGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 68
QY 94 TAAGCAATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 153
Db 69 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 128
QY 154 GCTTCAAAAAAGGAAA 169
Db 129 AAAAAAAATAAAAAAA 144
XX
RESULT 10
ABV04310/c
ID ABV04310 standard; cDNA; 300 BP.
XX
XX ABV04310;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 4301.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
```

```

PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Endege WO, Monahan JB;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 750; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 300 BP; 125 A; 22 C; 19 G; 101 T; 33 other;
XX
Query Match 9.0%; Score 47.8; DB 23; Length 300;
Best Local Similarity 56.2%; Pred. No. 0.068;
Matches 82; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 24 TGATTATATGCTGATTTATGGGTGATTTGCTCTCTTTATATCTTTATTTATCTCC 83
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 TCAAAAATTTNNTTTTTTTGGAAATTTNNTTTTTTTTCCTTTTAAATTTTTT 109
QY 84 AAATTTTCTTAAGCAATATTTCTTGCTAATCAATTAATATCAAGRAAARAAAC 143
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 TNNTTTTTAAATTAATTTTTTTTTTTTTCNCCAAAAAATAAAAAAAAAAAAAA 49
QY 144 TGAAGCAACGCTTCAAAAAAGGAA 169
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 AAAAAAGAAAAAATAAAAAAAAAAAAAA 23

RESULT 11
ABL34175/c
ID ABL34175 standard; DNA; 113515 BP.
XX
AC ABL34175;
XX
XX 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 2148.
XX
XX Human; immune system disease; cytosine methylation; antiaesthetic;
KW antihistaminic; antianemic; cytosine; cytosine; cytosine;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antidiabetic; antidiabetic; antipsoriasis;
KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.
XX

```

```

OS Homo sapiens.
XX
PN WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX Claim 1; SEQ ID NO 2148; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
XX Sequence 113515 BP; 35172 A; 1174 C; 22520 G; 54649 T; 0 other;
XX
Query Match 9.0%; Score 47.4; DB 24; Length 113515;
Best Local Similarity 54.4%; Pred. No. 0.49;
Matches 93; Conservative 1; Mismatches 77; Indels 0; Gaps 0;

QY 55 CTCTCTCTTTATCTTTTATTTATTTCCCAATTTTCTTAAGCAATATTTCTTTGCTA 114
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41662 CTCTCTCTTTATCAATTAATTAATTTCTTCTTATTTAAACAACTTTTCTTT 41603
QY 115 ATCAATAAATTTATCAAGAAAAAATACTGAAAGCAACGCTTGAAAAAGGAAAGTTAG 174
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41602 ATTTATATATTTACTATAAATACTATAAATTTCTAATAATTAATAAATTAATCA 41543
QY 175 CCCTATCGGGTATTTTGGAAAGTTGTAAATACTACTGTCTCTCTTAA 225
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41542 TTCCCACTTATTTATTTTAAATATCAATTTATCAAAATTTAACTTATA 41492

RESULT 12
ABV44994
ID ABV44994 standard; cDNA; 309 BP.
XX
XX AC ABV44994;
XX
XX 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 44985.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX

```

```

PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 8916; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 309 BP; 188 A; 13 C; 11 G; 97 T; 0 other;
XX
Query Match      8.9%; Score 47; DB 23; Length 309;
Best Local Similarity 58.0%; Pred. No. 0.11;
Matches 83; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 27 TTTATATGCTGATTATGGTGATTTTGGCTTCCTTCTTTATATCTTTTATTTATCCCAA 86
DB 33 TTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 92
QY 87 TTTTCTTAAGCAAAATTTCTTTGCTTAATAAATATCAAAAGAAAAAACTGA 146
DB 93 TTTTCTTTTAAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATTTG 152
QY 147 AAGCAACGCTTGAAGAAAGAAA 169
DB 153 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 175
RESULT 13
ABV4994/c
ID ABV4994 standard; cDNA; 309 BP.
XX
XX ABV4994;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 44985.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 8916; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 309 BP; 188 A; 13 C; 11 G; 97 T; 0 other;
XX
Query Match      8.9%; Score 47; DB 23; Length 309;
Best Local Similarity 58.0%; Pred. No. 0.11;
Matches 83; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 27 TTTATATGCTGATTATGGTGATTTTGGCTTCCTTCTTTATATCTTTTATTTATCCCAA 86
DB 207 TTTTCTTTTAAAGCAAAATTTCTTTGCTTAATAAATATCAAAAGAAAAAACTGA 146
QY 87 TTTTCTTTAAGCAAAATTTCTTTGCTTAATAAATATCAAAAGAAAAAACTGA 146
DB 147 TTTTCTTTTAAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATTTG 152
QY 147 AAGCAACGCTTGAAGAAAGAAA 169
DB 87 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 65
RESULT 14
ABV56492/c
ID ABV56492 standard; cDNA; 474 BP.
XX
XX ABV56492;
XX
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 56483.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 23-AUG-2001.
XX

```


PR 11-JAN-2000; 2000US-0480902.
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2; SEQ ID NO 8414; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=200201371139.
XX
XX Sequence 312 BP; 167 A; 27 C; 28 G; 90 T; 0 other;
XX
Query Match 8.8%; Score 46.6; DB 25; Length 312;
Best Local Similarity 57.0%; Pred. No. 0.13;
Matches 85; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 21 GGCTGATTTATGCTGATTGTTGGGTGATTTGCTTCCTTTATATCTTTATTTATT 80
Db 25 GCCTAATT 84
QY 81 CCCAAATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAA 140
Db 85 TTAATTTTTTTTTTTTTTTTATTAATAATAAAAAAATAAAAAAATAAAAAA 144
QY 141 AACTGAAGCAAGCTTTGAAAAAGGAAA 169
Db 145 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 173
RESULT 17
ABZ08201/c
ID ABZ08201 standard; cDNA; 480 BP.
XX
XX ABZ08201;
XX
XX 09-JAN-2003 (first entry)
DT Human leukocyte derived cDNA SEQ ID NO 8192.
DE

XX Human; leukocyte; gene expression profiling; allograft rejection;
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; gene;
KW ss.
XX
XX Homo sapiens.
OS
XX WO200257414-A2.
PN
XX 25-JUL-2002.
PD
XX 22-OCT-2001; 2001WO-US47856.
PF
XX 20-OCT-2000; 2000US-241994P.
PR
XX 08-JUN-2001; 2001US-296764P.
PR
XX (BIOC-) BIOCARDIA INC.
PA
XX Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
PI Ly N, Woodward R, Quettermous T, Johnson F;
PI
XX WPI; 2002-636525/68.
XX
XX New system for leukocyte expression profiling, diagnosing a disease, or
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT or congestive heart failure, comprises diagnostic oligonucleotides
XX
XX Claim 26; Page 1829-1830; 2038pp; English.
XX
XX The invention relates to a system for detecting gene expression, which
CC comprises one or two isolated DNA molecules that detect expression of a
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ0010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disease,
CC predicting therapeutic outcome, determining prognosis for a patient,
CC to treatment in an individual. The diseases include cardiac allograft
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection. The
CC present sequence is that of a human leukocyte expressed cDNA of the
CC invention.
XX
XX Sequence 480 BP; 124 A; 34 C; 159 G; 163 T; 0 other;
XX
Query Match 8.8%; Score 46.4; DB 24; Length 480;
Best Local Similarity 61.7%; Pred. No. 0.17;
Matches 74; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 50 TTTTGCTTCCTTCTTTTACTTTTATTTATTTCCCAATTTTCTTAAGCAATATTTCTT 109
Db 142 TTTTACTTTTTTTTTTCATATTTTTTAATACCAAAATAATAAAAAATAAAATATTTTA 83
QY 110 TGCTAATCAATAAATATCAAAAGAAAAAATAAAACCTGAAAGCAACGCTTCAAAAAAGGAAA 169
Db 82 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 23
RESULT 18
ABX43249/c
ID ABX43249 standard; cDNA; 312 BP.
XX
XX ABX43249;
XX
XX 20-FEB-2003 (first entry)
DT
XX Bovine EST associated with lactation/muscle/fat deposition #8414.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.

[illegible]

CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 10988 BP; 3128 A; 225 C; 2649 G; 4986 T; 0 other;
Query Match 8.4%; Score 44.6; DB 24; Length 10988;
Best Local Similarity 60.2%; Pred. No. 1.1;
Matches 74; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 12 TGTAGTCGTGGCTGATTTATATATGCTGATTTATGCTGATTTGCTCTCTTTATACCT 71
DB 10263 TGTAGATGCTGTTGATATATTTTGGTTTTATATTTTGGTTTTATAGTATACGTTT 10322
QY 72 TTATTATTCCTCAATTTTCTTAAGCAAAATATTTCTTGTCAATCAATAAATATCAAA 131
DB 10323 TTTTATATTTAAATGGATTTTGAGAAATTTATTTATTAGAAATTTAAAGTATATTCGAT 10382
QY 132 AGA 134
DB 10383 AAA 10385
RESULT 26
ID ABK31245 standard; DNA; 10988 BP.
XX
XX ABK31245;
AC
DT 23-APR-2002 (first entry)
XX
XX DE Signal transduction associated gene modified complementary DNA #44.
XX
XX Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
XX OS Homo sapiens.
XX Synthetic.
XX
XX WO200200926-A2.
XX
XX PD 03-JAN-2002.
XX
XX PF 29-JUN-2001; 2001WO-EP07472.
XX
XX PR 30-JUN-2000; 2000DE-1032529.
XX
XX PR 01-SEP-2000; 2000DE-1043826.
XX
XX (SPIG-) EPIGENOMICS AG.
XX
XX PA Olek A, Piepenbrock C, Berlin K;
XX
XX PI WPI; 2002-147896/19.
XX
XX DR Oligonucleotide for diagnosis and therapy of diseases associated with
XX signal transduction e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with signal transduction -
XX
XX PS Claim 1; SEQ ID No 88; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
XX signal transduction associated genes. The DNA sequences are chemically
XX modified using a solution of bisulphite, hydrogen sulphite or
XX disulphite. Also disclosed are oligonucleotides and/or DNA oligomers
XX for detecting the cytosine methylation state (CpG islands) of these
XX genes, and a method for the diagnosis and/or therapy of genetic and
XX epigenetic parameters of genes associated with signal transduction.
XX The genomic DNA can be obtained from cells or cellular components which
XX contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
XX cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
XX eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
XX histologic object slides, and all their possible combinations. The
XX sequences of the invention are useful for the diagnosis and therapy of
XX diseases associated with signal transduction e.g. solid tumours and
XX cancer. ABK3158-ABK31545 represent chemically pretreated genomic DNA
XX sequences of different genes associated with signal transduction, or
XX their complementary sequences.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX European Patent Office.
SQ Sequence 10988 BP; 3128 A; 225 C; 2649 G; 4986 T; 0 other;
Query Match 8.4%; Score 44.6; DB 24; Length 10988;
Best Local Similarity 60.2%; Pred. No. 1.1;
Matches 74; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 12 TGTAGTCGTGGCTGATTTATATATGCTGATTTATGCTGATTTGCTCTCTTTATACCT 71
DB 10263 TGTAGATGCTGTTGATATATTTTGGTTTTATATTTTGGTTTTATAGTATACGTTT 10322
QY 72 TTATTATTCCTCAATTTTCTTAAGCAAAATATTTCTTGTCAATCAATAAATATCAAA 131
DB 10323 TTTTATATTTAAATGGATTTTGAGAAATTTATTTATTAGAAATTTAAAGTATATTCGAT 10382
QY 132 AGA 134
DB 10383 AAA 10385
RESULT 27
ID AAS60450 standard; cDNA; 425 BP.
XX
XX AC AAS60450;
XX
XX DT 29-JAN-2002 (first entry)
XX
XX DE Human cancer agent-sensitive marker #181.
XX
XX OS Homo sapiens.
XX
XX KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmacytoma; reticulum cell sarcoma;
KW Hodgkin's disease; glioma; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200179556-A2.
XX
XX PD 25-OCT-2001.
XX
XX PF 13-APR-2001; 2001WO-US12132.
XX
XX PR 14-APR-2000; 2000US-197538P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX PI Lillie J, Brown JL, Bolt A, Van Huffel C;
XX
XX DR WPI; 2001-602933/68.
XX
XX PT Novel nucleic acid, used as a marker to determine the effectiveness of
XX using TAXOL to treat cancer cell growth in individuals -
XX
XX PS Claim 1; Page 221; 527pp; English.
XX
XX The invention relates to 1046 novel nucleic acids which are used as
XX markers for determining the sensitivity of a cancer cell to the
XX anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
XX they are shown to express one of the 242 sensitivity markers or the
XX cells are shown not to express one of the 804 resistance markers.
XX The methods can be used to determine the effectiveness of TAXOL
XX in the treatment of cancer cell growth in an individual. The markers
XX can be used as targets in developing anti-cancer agents such as
XX chemotherapeutic compounds. The markers can also be used as targets in
XX developing treatments for cancer, particularly those cancers which

CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 10988 BP; 3128 A; 225 C; 2649 G; 4986 T; 0 other;
Query Match 8.4%; Score 44.6; DB 24; Length 10988;
Best Local Similarity 60.2%; Pred. No. 1.1;
Matches 74; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 12 TGTAGTCGTGGCTGATTTATATATGCTGATTTATGCTGATTTGCTCTCTTTATACCT 71
DB 10263 TGTAGATGCTGTTGATATATTTTGGTTTTATATTTTGTGTTTATAGTATACGTTT 10322
QY 72 TTATTATTCCTCAATTTTCTTAAGCAAAATATTTCTTGTCAATCAATAAATATCAAA 131
DB 10323 TTTTATATTTAAATGGATTTGAGAAATTTATTTATTAGAAATTTAAAGTATATTCGAT 10382
QY 132 AGA 134
DB 10383 AAA 10385
RESULT 26
ID ABK31245 standard; DNA; 10988 BP.
XX
XX ABK31245;
AC
DT 23-APR-2002 (first entry)
XX
XX DE Signal transduction associated gene modified complementary DNA #44.
XX
XX Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
XX OS Homo sapiens.
XX Synthetic.
XX
XX WO200200926-A2.
XX
XX PD 03-JAN-2002.
XX
XX PF 29-JUN-2001; 2001WO-EP07472.
XX
XX PR 30-JUN-2000; 2000DE-1032529.
XX
XX PR 01-SEP-2000; 2000DE-1043826.
XX
XX PA (SPIG-) EPIGENOMICS AG.
XX
XX PI Olek A, Piepenbrock C, Berlin K;
XX
XX DR WPI; 2002-147896/19.
XX
XX PT Oligonucleotide for diagnosis and therapy of diseases associated with
XX signal transduction e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with signal transduction -
XX
XX PS Claim 1; SEQ ID No 88; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
XX signal transduction associated genes. The DNA sequences are chemically
XX modified using a solution of bisulphite, hydrogen sulphite or
XX disulphite. Also disclosed are oligonucleotides and/or DNA oligomers
XX for detecting the cytosine methylation state (CpG islands) of these
XX genes, and a method for the diagnosis and/or therapy of genetic and
XX epigenetic parameters of genes associated with signal transduction.
XX The genomic DNA can be obtained from cells or cellular components which
XX contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
XX cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
XX eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
XX histologic object slides, and all their possible combinations. The
XX sequences of the invention are useful for the diagnosis and therapy of
XX diseases associated with signal transduction e.g. solid tumours and
XX cancer. ABK3158-ABK31545 represent chemically pretreated genomic DNA
XX sequences of different genes associated with signal transduction, or
XX their complementary sequences.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX European Patent Office.
SQ Sequence 10988 BP; 3128 A; 225 C; 2649 G; 4986 T; 0 other;
Query Match 8.4%; Score 44.6; DB 24; Length 10988;
Best Local Similarity 60.2%; Pred. No. 1.1;
Matches 74; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 12 TGTAGTCGTGGCTGATTTATATATGCTGATTTATGCTGATTTGCTCTCTTTATACCT 71
DB 10263 TGTAGATGCTGTTGATATATTTTGGTTTTATATTTTGTGTTTATAGTATACGTTT 10322
QY 72 TTATTATTCCTCAATTTTCTTAAGCAAAATATTTCTTGTCAATCAATAAATATCAAA 131
DB 10323 TTTTATATTTAAATGGATTTGAGAAATTTATTTATTAGAAATTTAAAGTATATTCGAT 10382
QY 132 AGA 134
DB 10383 AAA 10385
RESULT 27
ID AAS60450 standard; cDNA; 425 BP.
XX
XX AC AAS60450;
XX
XX DT 29-JAN-2002 (first entry)
XX
XX DE Human cancer agent-sensitive marker #181.
XX
XX OS Homo sapiens.
XX
XX KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmacytoma; reticulum cell sarcoma;
KW Hodgkin's disease; glioma; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200179556-A2.
XX
XX PD 25-OCT-2001.
XX
XX PF 13-APR-2001; 2001WO-US12132.
XX
XX PR 14-APR-2000; 2000US-197538P.
XX
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX PI Lillie J, Brown JL, Bolt A, Van Huffel C;
XX
XX DR WPI; 2001-602933/68.
XX
XX PT Novel nucleic acid, used as a marker to determine the effectiveness of
XX using TAXOL to treat cancer cell growth in individuals -
XX
XX PS Claim 1; Page 221; 527pp; English.
XX
XX The invention relates to 1046 novel nucleic acids which are used as
XX markers for determining the sensitivity of a cancer cell to the
XX anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
XX they are shown to express one of the 242 sensitivity markers or the
XX cells are shown not to express one of the 804 resistance markers.
XX The methods can be used to determine the effectiveness of TAXOL
XX in the treatment of cancer cell growth in an individual. The markers
XX can be used as targets in developing anti-cancer agents such as
XX chemotherapeutic compounds. The markers can also be used as targets in
XX developing treatments for cancer, particularly those cancers which

KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; anti anaemic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX Homo sapiens.
 OS
 XX WO200200928-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP07537.
 PF
 XX 30-JUN-2000; 2000DE-1032529.
 PR
 XX 01-SEP-2000; 2000DE-1043826.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 PI
 XX WPI; 2002-130909/17.
 DR
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 XX for diagnosis and treatment of diseases associated with abnormal
 XX cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 1110; 32pp + Sequence Listing; German.
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 XX Sequence 6078 BP; 1615 A; 218 C; 1499 G; 2746 T; 0 other;
 SQ
 Query Match 8.4%; Score 44.4; DB 24; Length 6078;
 Best Local Similarity 54.4%; Pred. No. 1.1;
 Matches 87; Conservative 1; Mismatches 72; Indels 0; Gaps 0;
 QY 65 TATACCTTTTATTTCCCAAAATTTTCTTAAGCAAAATATTTCTTGTAAATCAATAAAT 124
 DB 4494 TAAATATATACCAAAATCTTAATATATATAAAAAAACCCTCACACATAAATAAAA 4435
 QY 125 TATCAAGAGAAAAAAGAACTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATCGGG 184
 DB 4434 AAAAAAAGAAAAAACAATATTAACCAAACTCTTCAATATATATCTTAAATACCTTTT 4375
 QY 185 TATATTTTGAAGTGTGAAATACTACRTGTTCTCTCTTA 224
 DB 4374 TATTTTATATTTTATATATTTTCCAAATTTCTACAA 4335
 RESULT 32
 ID AAS07727
 XX AAS07727 standard; DNA; 198 BP.
 AC AAS07727;
 XX
 XX 23-OCT-2001 (first entry)
 DE
 DE Cervical cancer pre-malignant condition DNA marker #26.
 KW Cervical cancer; pre-malignant condition marker; tumorigenesis; CIN; ds;
 KW cervical intraepithelial neoplasia; squamous intraepithelial lesions;
 KW polymerase chain reaction; PCR; enzyme linked immunosorbent assay; ELISA;

KW anticancer therapy; carcinogen; antisense inhibition.
 XX Homo sapiens.
 XX WO200142792-A2.
 PN
 XX 14-JUN-2001.
 PD
 XX 08-DEC-2000; 2000WO-US33311.
 PF
 XX 08-DEC-1999; 99US-0169811.
 PR
 XX 21-DEC-1999; 99US-0171330.
 PR
 XX 31-MAR-2000; 2000US-0189113.
 PR
 XX 31-MAR-2000; 2000US-0193943.
 PR
 XX 12-MAY-2000; 2000US-0203772.
 PR
 XX 09-JUN-2000; 2000US-0210820.
 PR
 XX 21-JUL-2000; 2000US-0220113.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI
 XX Schlegel R, Deeds J, Berger A, Zhao X;
 PI
 XX WPI; 2001-367889/38.
 DR
 XX Cervical cancer protein markers useful for the diagnosis, prevention
 XX and treatment of cervical cancers, especially cervical intraepithelial
 XX neoplasia or squamous intraepithelial lesions -
 XX
 PS 20; Page 407; 436pp; English.
 CC The sequence represents a cervical cancer pre-malignant condition marker.
 CC These markers encode proteins that are over expressed during
 CC tumorigenesis. The proteins and their corresponding nucleic acid
 CC sequences can therefore be used for the diagnosis, prevention and
 CC treatment of cervical cancers, particularly cervical intraepithelial
 CC neoplasia (CIN) and squamous intraepithelial lesions (SIL). The sequences
 CC may be used as markers in diagnostic assays to detect cancerous
 CC conditions (e.g. by polymerase chain reaction (PCR) or enzyme linked
 CC immunosorbent assay (ELISA)), to monitor the efficacy of anticancer
 CC therapies and to identify anticancer or carcinogenic compounds. The level
 CC of expression of a marker in a patient sample is compared with the normal
 CC level of expression of the marker in a control non-cervical cancer
 CC sample, whereby a significant difference indicates that the patient is
 CC afflicted with cervical cancer or a pre-malignant condition. The
 CC expression of these proteins may be inhibited by antisense inhibition for
 CC the treatment of cancers. They may be used in this way for the treatment
 CC of cervical intraepithelial neoplasia or squamous intraepithelial
 XX lesions.
 XX
 SQ Sequence 198 BP; 67 A; 11 C; 10 G; 102 T; 8 other;
 Query Match 8.3%; Score 43.8; DB 22; Length 198;
 Best Local Similarity 54.5%; Pred. No. 0.53;
 Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 QY 27 TTTATATGCTGATTTATGGTGATTTTCTCTCTTCTTATCTTTTATTTATCCCAA 86
 DB 40 TTTTCTCTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 99
 QY 87 TTTTCTTAAAGCAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACTGA 146
 DB 100 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 159
 QY 147 RAGCAACGCTTGAAAAAGGAAA 169
 DB 160 AANANAAAAAANAAAAAANAAA 182
 RESULT 33
 ID ABX47018
 XX ABX47018 standard; cDNA; 236 BP.
 AC ABX47018;


```
XX PN WO200200928-A2.
XX PD 03-JAN-2002.
XX PF 02-JUL-2001; 2001WO-EP07537.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-130909/17.
XX PT Nucleic acid comprising fragment of chemically modified gene, useful
XX PT for diagnosis and treatment of diseases associated with abnormal
XX PT cytosine methylation -
XX PS Claim 1; SEQ ID NO 1449; 32pp + Sequence Listing; German.
XX CC The present invention provides a number of human immune system associated
XX CC genes which are modified by the methylation of cytosines. The sequences
XX CC can be used in the diagnosis and treatment of immune system disorders,
XX CC including eye diseases such as retinopathy, neovascular glaucoma and
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX CC diseases. The present sequence is a gene of the invention.
XX SQ Sequence 6590 BP; 2030 A; 155 C; 1534 G; 2871 T; 0 other;

Query Match      8.2%; Score 43.6; DB 24; Length 6590;
Best Local Similarity 57.2%; Pred. No. 1.7;
Matches 79; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTTATGGGCTGATTTGCTTCCTCTTTTATATCTTTTATTTATTTCCCAA 86
DB 1088 TTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1029

QY 87 TTTTCTTAAGCAATATTTCTTTGCTAATCAATTAATCAAAAGAAAAAACTGA 146
DB 1028 TATTTCTCGAAAAAAATATATAAATCAATAAATCAATAAATAAATAAATAAATCAACA 969

QY 147 AAGCAAGCTTGAAAAA 164
DB 968 AATAAACAGGTAAACAAA 951

RESULT 41
AAH70049
ID AAH70049 standard; cDNA; 300 BP.
XX AC AAH70049;
XX DT 19-SEP-2001 (first entry)
XX DE Human cervical cancer marker nucleic acid 1323.
XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX OS Homo sapiens.
XX PN WO200142467-A2.
XX PD 14-JUN-2001.
XX PF 08-DEC-2000; 2000WO-US33312.
XX PR 08-DEC-1999; 99US-0169681.
XX PR 21-DEC-1999; 99US-0171350.
XX PR 14-MAR-2000; 2000US-0189315.
```

```
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Deeds J, Berger A, Zhao X;
XX DR WPI; 2001-375006/39.
XX PS New isolated nucleic acid for diagnosing and treating cervical cancer
XX PT and for assessing and detecting compounds for treating the cancer -
XX CC Claim 1; Page 309; 1051pp; English.
XX CC The invention relates to novel genes (AAH68727-AAH71383) associated with
XX CC cervical cancer with cytostatic activity. The nucleic acids and encoded
XX CC polypeptides are useful: to assess if a patient is afflicted with
XX CC cervical cancer or has a pre-malignant condition; to monitor the
XX CC progression of cervical cancer or a premalignant condition in a patient;
XX CC and to select and/or assess the efficacy of a compound or therapy for
XX CC inhibiting cervical cancer in a patient. The nucleic acids may also be
XX CC useful for gene therapy.
XX SQ Sequence 300 BP; 81 A; 27 C; 20 G; 143 T; 29 other;

Query Match      8.2%; Score 43.4; DB 22; Length 300;
Best Local Similarity 48.0%; Pred. No. 0.75;
Matches 98; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 4 GGTGATGCTGATGCTGCTGCTGATTTATATGCTGATTTATGGGTGATTTGCTTCCTTCT 63
DB 25 GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 84

QY 64 TTATACTTTTATTTATTTCCCAATTTTCTTAAGCAATATTTCTTCTGCTAATCAATAA 123
DB 85 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAA 144

QY 124 TTATCAAAAGAAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATCGG 183
DB 145 NAAAAAANNNAANTNANGGNGNNAAAAAANTTTTNAAAAAAANTTTTNCCTAATTNGG 204

QY 184 GTATATTTTGGAAGTGTGAAAAATA 207
DB 205 GTTTTAAAGGGAAGAAAAA 228

RESULT 42
ABX46053
ID ABX46053 standard; cDNA; 424 BP.
XX AC ABX46053;
XX DT 21-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #11218.
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
XX KW muscle deposition; fat deposition; genome mapping; gene identification;
XX KW gene analysis; cattle breeding.
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-0960352.
XX PR 12-JAN-1999; 99US-115707P.
XX PR 11-JAN-2000; 2000US-0480902.
XX PA (BYAT/) BYATT J C.
```


DT XX 21-FEB-2003 (first entry)

DE XX Bovine EST associated with lactation/muscle/fat deposition #9335.

KW XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

KW XX muscle deposition; fat deposition; genome mapping; gene identification;

KW XX gene analysis; cattle breeding.

OS XX Bos Taurus.

XX XX

PN XX US2002137139-A1.

XX XX

PD XX 26-SEP-2002.

XX XX

PF XX 24-SEP-2001; 2001US-0960352.

XX XX

PR XX 12-JAN-1999; 99US-115707P.

PR XX 11-JAN-2000; 2000US-0480902.

XX XX

PA (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

XX XX

PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX XX

DR WPI; 2003-110599/10.

XX XX

PT New nucleic acid associated with lactation, and muscle and fat

PT deposition, useful for genome mapping, gene identification and

PT analysis, cattle breeding, or for genetically improving cattle

XX XX

PS Claim 2; SEQ ID No 9335; 245pp; English.

XX XX

CC The invention relates to a purified nucleic acid molecule associated with

CC lactation or muscle and fat deposition (designated LMFD), derived

CC from cattle, and the LMFD nucleic acid can specifically hybridize to a

CC second nucleic acid molecule comprising any of 15112 nucleotide

CC sequences, appearing as ABX34836-ABX49947, or complements of them.

CC Also included are: (1) a transformed cell having a nucleic acid

CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-

CC translated sequence that functions in the cell to cause termination of

CC transcription and addition of polyadenylated ribonucleotides to a 3' end

CC of the mRNA molecule; and (2) determining a level or pattern of a

CC nucleic acid in a bovine cell or tissue comprising: (a) incubating a marker

CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its

CC complement or fragment) with a complementary nucleic acid molecule

CC obtained from the bovine cell or tissue, where hybridisation between the

CC marker nucleic acid and the complementary nucleic acid permits the

CC detection of the molecule; and (b) detecting the level or pattern of the

CC complementary nucleic acid, where the detection of the complementary

CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMFD nucleic acid is used for determining a level or pattern

CC of a molecule in a bovine cell or tissue. It is useful for genome

CC mapping, gene identification and analysis, cattle breeding, preparation

CC of constructs for use in cattle gene expression, or for genetically

CC improving cattle. The present sequence is one of the 15112 bovine

CC LMFD EST (expressed sequence tag) nucleic acids.

CC Note: The present sequence was not shown in the specification but

CC was obtained in electronic format from the USPTO web site:

CC seqdata.uspto.gov/sequence.html?DocID=20020137139.

XX XX

SQ Sequence 380 BP; 154 A; 22 C; 46 G; 158 T; 0 other;

Query Match 8.2%; Score 43.2; DB 25; Length 380;

Best Local Similarity 56.2%; Pred. No. 0.9;

Matches 81; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTTATGGTGATTTTCTCTCTTCTTATCTTTTATTTATCCCAA 86

DB 49 TTTTGGTTGGTT 108

QY 87 TTTTCTTAAGCAAAATTTCTTCTGCTAATCAATAAATATCAAAAGAAAAAACTGA 146

DB 109 TTTTCTTATTTATAATAATTTTCAAAATTTAAAAAATAATAATTTGTTTAAAAAATTTT 168

QY 147 AAGCAACGCTTGAAGAAAGGAAG 170

DB 169 AAAAAAATTTAAAAAAGAAAAAG 192

RESULT 45

ABV83644

ID ABV83644 standard; cDNA; 430 BP.

XX XX

AC ABV83644;

XX XX

DT 06-DEC-2002 (first entry)

XX XX

DE Human breast specific gene SEQ ID NO 87.

XX XX

KW Human; breast; cytostatic; cancer; transgenic; gene therapy; vaccine;

KW KW gene; ss.

XX XX

OS Homo sapiens.

XX XX

PN WO200266605-A2.

XX XX

PD 29-AUG-2002.

XX XX

PF 14-FEB-2002; 2002WO-US04284.

XX XX

PR 15-FEB-2001; 2001US-268999P.

XX XX

PA (DIAD-) DIADEXUS INC.

XX XX

PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;

PI Sun Y, Liu C;

XX XX

DR WPI; 2002-713345/77.

XX XX

PT New isolated breast specific nucleic acid molecules and polypeptides,

PT useful for identifying, diagnosing, monitoring, staging, imaging and

PT treating breast cancer and non-cancerous disease states in breast

PT tissue

XX XX

PS Claim 1; Page 193-194; 254pp; English.

XX XX

CC The invention relates to human breast specific nucleic acids (I)

CC comprising:

CC (a) a sequence encoding any one of 95 protein sequences

CC (ABP66614-ABP66708);

CC (b) any one of 115 polynucleotide sequences (ABV83558-ABV83672);

CC (c) a molecule that selectively hybridizes to (a) or (b);

CC (d) a molecule having at least 60% sequence identity to (a) or (b).

CC The breast specific nucleic acid molecules, polypeptides and antibodies

CC are useful for identifying, diagnosing, monitoring, staging, imaging and

CC treating breast cancer and non-cancerous disease states in breast tissue.

CC They are also useful for producing transgenic animals and cells and

CC producing engineered breast tissue for treatment and research. The

CC transgenic animals are useful as animal model systems used in elaborating

CC the biological function of the polypeptides, studying conditions and/or

CC disorders associated with aberrant expression and in screening for

CC compounds effective in ameliorating the conditions. The polynucleotides

CC are useful for gene therapy and in vaccines.

XX XX

SQ Sequence 430 BP; 135 A; 97 C; 103 G; 95 T; 0 other;

Query Match 8.2%; Score 43.2; DB 24; Length 430;

Best Local Similarity 60.0%; Pred. No. 0.93;

Matches 72; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 50 TTTTGTCTCTCTTATCTTATCTTTATTTATTTCCCAAAATTTTCTTAAGCAAAATTTCTT 109

DB 52 TTTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 111

Qy 110 TGCTAATCAATAAATTATCAAAAGAAAAAACTGAAAGCAACGCTTCGAAAAAGGAAA 169
Db 112 TTAAAAAGGGAATTTATAAAAAAAAGGTTTGACCCCAAAAAATTTAAA 171

Search completed: February 14, 2004, 14:10:12
Job time : 285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 13:57:05 ; Search time 71 Seconds

(without alignments)

3288.615 Million cell updates/sec

Title: US-10-027-632-1

Perfect score: 529

Sequence: 1 catggatgctgtgctgt.....tcttgagaggcttctgag 529

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	49.2	9.3	5852	1	US-07-867-106-2
C 2	44.4	8.4	664	4	US-09-904-615-66
C 3	42.4	8.0	2672	1	US-08-703-947-1
C 4	42	7.9	1342	4	US-09-489-847-89
C 5	40	7.6	1544	4	US-09-187-999-14
C 6	39.8	7.5	8920	2	US-08-446-855A-1
C 7	39.8	7.5	8920	3	US-09-150-741-1
C 8	39.6	7.5	664	4	US-09-904-615-66
C 9	39.4	7.4	873	3	US-09-475-316A-20
C 10	38.8	7.3	1378	4	US-09-149-476-208
C 11	38.8	7.3	2323	4	US-09-149-476-24
C 12	38.8	7.3	5232	3	US-09-212-971-3
C 13	38.8	7.3	5232	3	US-08-800-925A-3
C 14	38.8	7.3	5232	4	US-09-617-053A-3
C 15	38	7.2	1582	3	US-08-545-196B-10
C 16	38	7.2	1582	3	US-08-545-196B-12
C 17	38	7.2	18953	4	US-09-820-005-3
C 18	37.8	7.1	5656	4	US-09-410-464-5
C 19	37.6	7.1	732	4	US-09-149-476-66
C 20	37.6	7.1	746	3	US-09-013-810-1
C 21	37.6	7.1	3028	4	US-09-221-013A-7
C 22	37.6	7.1	4344	4	US-09-660-198-165
C 23	37.6	7.1	1664976	4	US-08-916-421B-1
C 24	37.2	7.0	1492	4	US-09-369-247-23
C 25	37.2	7.0	5852	1	US-07-867-106-2
C 26	37	7.0	1582	3	US-08-545-196B-10
C 27	37	7.0	1582	3	US-08-545-196B-12

28	37	7.0	2007	3	US-08-747-221B-36	Sequence 36, Appl
C 29	37	7.0	2007	3	US-08-747-221B-38	Sequence 38, Appl
30	37	7.0	2007	3	US-09-005-051-36	Sequence 36, Appl
C 31	37	7.0	2007	3	US-09-005-051-38	Sequence 38, Appl
32	37	7.0	2334	1	US-08-062-632-4	Sequence 4, Appl
33	37	7.0	3715	4	US-09-234-245-1	Sequence 1, Appl
34	36.8	7.0	16442	3	US-08-781-891-208	Sequence 208, App
35	36.8	7.0	16442	3	US-09-618-166-208	Sequence 208, App
C 36	36.4	6.9	1037	4	US-09-489-847-112	Sequence 112, App
37	36.4	6.9	4140	3	US-08-894-731-2	Sequence 2, Appl
38	36.2	6.8	318	4	US-09-328-352-642	Sequence 642, App
39	36	6.8	1507	4	US-09-453-323-1	Sequence 1, Appl
40	36	6.8	1577	3	US-08-821-994-59	Sequence 59, Appl
41	36	6.8	2409	3	US-09-293-322C-8	Sequence 8, Appl
42	36	6.8	2409	3	US-09-839-497A-8	Sequence 8, Appl
C 43	36	6.8	3165	4	US-09-601-198-102	Sequence 102, App
C 44	36	6.8	3581	2	US-08-738-349-1	Sequence 1, Appl
C 45	35.6	6.7	519	4	US-09-227-357-76	Sequence 76, Appl
C 46	35.6	6.7	1782	4	US-09-134-001C-1154	Sequence 1154, Ap
47	35.6	6.7	3636	3	US-09-074-579-2	Sequence 2, Appl
48	35.6	6.7	3636	3	US-09-388-774-2	Sequence 2, Appl
49	35.4	6.7	1540	3	US-08-977-001-2	Sequence 2, Appl
C 50	35.4	6.7	1701	4	US-09-996-243-114	Sequence 114, App
C 51	35.4	6.7	1864	4	US-09-149-476-130	Sequence 130, App
C 52	35.4	6.7	2550	6	5258287-23	Patent No. 5258287
53	35.4	6.7	3214	1	US-08-484-105-17	Sequence 17, Appl
54	35.4	6.7	3214	1	US-08-484-106-17	Sequence 17, Appl
55	35.4	6.7	3232	4	US-09-333-214-3	Sequence 3, Appl
56	35.4	6.7	5727	4	US-09-628-188A-1	Sequence 1, Appl
C 57	35.4	6.7	11469	4	US-09-367-895-29	Sequence 29, Appl
C 58	35.4	6.7	11478	3	US-08-981-803-29	Sequence 29, Appl
C 59	35.4	6.7	11478	3	US-08-983-440-29	Sequence 29, Appl
C 60	35.2	6.7	593	4	US-09-904-615-59	Sequence 59, Appl
C 61	35.2	6.7	599	3	US-09-328-111-147	Sequence 147, App
C 62	35.2	6.7	1172	1	US-07-945-288-9	Sequence 9, Appl
C 63	35.2	6.7	1172	1	US-08-462-831-9	Sequence 9, Appl
64	35.2	6.7	1172	1	US-08-461-809-9	Sequence 9, Appl
65	35.2	6.7	1172	1	US-08-461-441-9	Sequence 9, Appl
66	35.2	6.7	1172	5	PCT-US93-08518-9	Sequence 9, Appl
67	35.2	6.7	9951	4	US-09-193-707-3	Sequence 3, Appl
68	35.2	6.7	10524	4	US-09-193-707-4	Sequence 4, Appl
69	35.2	6.7	11927	4	US-09-193-707-5	Sequence 5, Appl
70	35	6.6	1248	4	US-09-489-847-101	Sequence 101, App
71	35	6.6	1466	4	US-08-984-919A-10	Sequence 10, Appl
C 72	35	6.6	1466	4	US-08-984-919A-12	Sequence 12, Appl
C 73	35	6.6	1472	3	US-08-781-420-10	Sequence 10, Appl
C 74	35	6.6	1472	3	US-08-781-420-12	Sequence 12, Appl
C 75	35	6.6	1472	4	US-08-874-102-10	Sequence 10, Appl
C 76	35	6.6	1472	4	US-08-874-102-12	Sequence 12, Appl
C 77	35	6.6	1472	4	US-09-006-595A-10	Sequence 10, Appl
C 78	35	6.6	1472	4	US-09-006-595A-12	Sequence 12, Appl
C 79	35	6.6	1474	3	US-08-821-994-64	Sequence 64, Appl
C 80	35	6.6	1875	4	US-08-984-919A-46	Sequence 46, Appl
C 81	35	6.6	1875	4	US-08-984-919A-48	Sequence 48, Appl
C 82	35	6.6	1881	4	US-08-874-102-46	Sequence 46, Appl
C 83	35	6.6	1881	4	US-08-874-102-48	Sequence 48, Appl
C 84	35	6.6	4055	4	US-09-620-312D-706	Sequence 706, App
C 85	35	6.6	6718	2	US-08-962-284-1	Sequence 1, Appl
C 86	35	6.6	168575	4	US-09-426-250-1	Sequence 1, Appl
C 87	34.8	6.6	3253	4	US-09-759-359A-1	Sequence 1, Appl
C 88	34.8	6.6	4117	4	US-09-484-970B-2	Sequence 2, Appl
C 89	34.6	6.5	469	4	US-09-229-583A-18	Sequence 18, Appl
C 90	34.6	6.5	1313	4	US-09-149-476-112	Sequence 112, App
C 91	34.6	6.5	1332	2	US-09-057-762-1	Sequence 1, Appl
C 92	34.6	6.5	1332	3	US-08-326-119A-1	Sequence 1, Appl
C 93	34.6	6.5	1342	4	US-09-489-847-89	Sequence 89, Appl
C 94	34.6	6.5	1637	4	US-09-205-258-178	Sequence 178, App
C 95	34.6	6.5	1827	4	US-09-134-001C-1914	Sequence 1914, Ap
C 96	34.6	6.5	3701	3	US-08-845-258-10	Sequence 10, Appl
C 97	34.6	6.5	3701	3	US-08-990-571-10	Sequence 10, Appl
C 98	34.6	6.5	3701	3	US-08-723-142A-10	Sequence 10, Appl
C 99	34.6	6.5	3701	4	US-09-528-784A-10	Sequence 10, Appl
C 100	34.6	6.5	3701	4	US-09-569-098A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Sline Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
US-07-867-106-2

Query Match 9.3%; Score 49.2; DB 1; Length 5852;
Best Local Similarity 57.0%; Pred. No. 0.00077;
Matches 90; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 12 TGTAGTCGGTGCTGATTATATGCTGATTATGGGATTTTGGCTTCCTCTTTTACTTT 71
Db 5767 TTTAATATTGTTATGTTATTTTATATATGTTATTTGTTGTTGTTTACTT 5708
Qy 72 TTATTATTCCTCCAAATTTTCTTAAGCAATATTTCTTGCTAATCAATAAATATCAAA 131
Db 5707 ATATTCTATTTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 5648
Qy 132 AGAAAAAAGCTGAAGCAAGCTGAAAAAGGAA 169
Db 5647 AAAAAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5610

RESULT 2

US-09-904-615-66/c
; Sequence 66, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032PI
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-66

Query Match 8.4%; Score 44.4; DB 4; Length 664;
Best Local Similarity 52.8%; Pred. No. 0.006;
Matches 93; Conservative 1; Mismatches 82; Indels 0; Gaps 0;

Qy 27 TTTATATGCTGATTATGCTGATTGCTTCCTCTTTTACTTTTATTTATCCCAA 86
Db 641 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 582
Qy 87 TTTTCTTTTAAAGCAATATTTCTTGCTAATCAATAAATATCAAAAGAAAAAACTGA 146
Db 581 ATTTCTGGAAGTTAAAGTAGATACAGCAATATACCAAAAAAAGAAAAAAGACA 522
Qy 147 AAGCAACCTTGAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGAAGTTGTA 202
Db 521 AAAAACTTCAATAATAATAATTTTACACTATGAAGTACACATTTGGAATTTGAA 466

RESULT 3

US-08-703-947-1/c
; Sequence 1, Application US/08703947
; Patent No. 5788962
; GENERAL INFORMATION:

APPLICANT: Wise, Kim S.
APPLICANT: McIntosh, Mark A.
TITLE OF INVENTION: DNA Sequences Coding for Mycoplasma
TITLE OF INVENTION: Hyopneumoniae Surface Antigens,
TITLE OF INVENTION: Corresponding Proteins and Use in
TITLE OF INVENTION: Vaccines and Diagnostic Procedures
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace J. Fishel
STREET: 929 Fee Fee Road, Suite 100
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63043
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 KB storage
; COMPUTER: Hewlett-Packard Vectra
; OPERATING SYSTEM: MS-DOS Version 3.3

SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/703,947
FILING DATE: 28-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/373,957
FILING DATE: January 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Fishel, Grace J.
REGISTRATION NUMBER: 25864
REFERENCE/DOCKET NUMBER: UYM 8141
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 878-0440
TELEFAX: (314) 275-7693
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2672 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
DESCRIPTION: region of 5.8 kb HindIII fragment from
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Mycoplasma hyopneumoniae
STRAIN: J
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE: unicellular bacterium
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY: Genomic in Charon 4A, GEM12
CLONE: MhpJ25, MhpJ35, MhpJG35, pZJ25, pZJ25.1, pZJ25.14,
pZJG35.1, pZJG35.12, pZJG35.13, pZJG35.14
POSITION IN GENOME:
CHROMOSOME/SEGMENT: single chromosome
MAP POSITION: unknown
UNITS: unknown
FEATURE:
NAME/KEY: sequence encodes entire 627 amino acids of
the structural gene for the surface lipoprotein
NAME/KEY: P65 and includes 312 bp upstream and 479
bp downstream of coding sequence
LOCATION: coding sequence for P65 spans 1881 bp of
sequence (begins at nt 313 and
ends at nt 2193)
LOCATION: includes all sequence through nt 2193
IDENTIFICATION METHOD: by similarity to pattern of open reading
frame; by experiment identifying protein products of
immunogenetic region of Mycoplasma
IDENTIFICATION METHOD: sequence with immune serum to P65
OTHER INFORMATION: immunogenic surface lipoprotein of no known
function; C-terminus exposed on external
surface of cell
PUBLICATION INFORMATION:
AUTHORS: Mary F. Kim, Manijeh B. Heidari, Susan J.
Stull, Mark A. McIntosh, and Kim S. Wise
TITLE: Identification and Mapping of an
Immunogenic Region of Mycoplasma
hyopneumoniae p65 Surface Lipoprotein
TITLE: Expressed in Escherichia coli from a Cloned
Genomic Fragment
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 8
PAGES: 2637-2643
DATE: August 1990

DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: From 1 to 2672
US-08-703-947-1
Query Match 8.0%; Score 42.4; DB 1; Length 2672;
Best Local Similarity 53.7%; Pred. No. 0.04;
Matches 88; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
Qy 40 TTATGGGTGATTTTGGCTTCTTTATACCTTTTATTTATTTCCCAATTTTCTTAAGCA 99
Db 340 TTAAGAAATTTTCTTGGCTTTTCTTCAATTTTCTTCTATATTTTGTATAGATGG 281
Qy 100 AATATTTCTTGTCAATTAATTAATCAAGAAAAAACTGAAAGCAAGCGCTTGA 159
Db 280 ACTAAATTTTACCATATTCCTAAAGAAAAAAAGTTTAAAGTTTAAATATAT 221
Qy 160 AAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAA 203
Db 220 AAAATCAACAAGATTTTATATTTTCTAAATTAATTTCAAAA 177
RESULT 4
US-09-489-847-89/c
Sequence 89, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
EARLIER FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1993-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 89
LENGTH: 1342
TYPE: DNA
ORGANISM: Homo sapiens
US-09-489-847-89
Query Match 7.9%; Score 42; DB 4; Length 1342;
Best Local Similarity 51.1%; Pred. No. 0.037;
Matches 96; Conservative 1; Mismatches 91; Indels 0; Gaps 0;
Qy 27 TTTATATGCTGATTTATGGGTGATTTTGGCTTCTTATACCTTTTATTTTATTTTCCCAA 86
Db 1341 TTTTCTTAAAGCAATATTTCTTTGCTAAATCAATTAATTAATCAAGAAAAAACTGA 146
Qy 87 TTTTCTTAAAGCAATATTTCTTTGCTAAATCAATTAATTAATCAAGAAAAAACTGA 146
Db 1281 TTTTCTTAAAGCAATATTTCTTTGCTAAATTAATTAATTAATCAAGAAAAAACTGA 1222
Qy 147 AAGCAAGCTTTGAAAGGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAAAT 206
Db 1221 TCTTAAACATTTACAAATTTGTACAAAGATTTGTAGCTTTTATATTTTAAAGATGCTAT 1162
Qy 207 ACTACRTG 214
Db 1161 ACTAAGAG 1154

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 8920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic
US-08-446-855A-1

Query Match          7.5%; Score 39.8; DB 2; Length 8920;
Best Local Similarity 49.7%; Pred. No. 0.35;
Matches 98; Conservative 1; Mismatches 98; Indels 0; Gaps 0

QY      29 TATATGCTGATTTAAGGGTGAATTTTGCTTCCTTTATACTTTTTATTTATCCCAAATT 88
Db      531 TATAAATTTATTTATATATATTTATTAATTTATTCAATTTATTTATTTTCTTAGT 590

QY      89 TTTCCTTAAGCAAAATATTTCTTTGCTTAATCAATAAATTTATCAAAGAAGAAAAAACTGAAA 148
Db      591 TTATAAATAGTAATTTCTACTAAATTTAAAAAAGAAAAAAGAAAAAAGAAAA 650

QY      149 GCAACGCTTGAAAAAGGAAAGTTAGGCCCTATCGGGTATATTTTGGAAAGTTGTAAAAATAC 208
Db      651 AAAAAAATTTACATATGAAAAATGAACCTGTATATCTAAATTTATAAATATTTTAAACA 710

QY      209 TACRTGTTCTCTTTCTAA 225
Db      711 TAAATATAAATGTATAA 727


RESULT 7
US-09-150-741-1
; Sequence 1, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; Patent No. 6183996
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PL6380
; EARLIER FILING DATE: 1992-12-16
; EARLIER APPLICATION NUMBER: AU93/00617
; EARLIER FILING DATE: 1993-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8920
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
US-09-150-741-1

Query Match          7.5%; Score 39.8; DB 3; Length 8920;
Best Local Similarity 49.7%; Pred. No. 0.35;
Matches 98; Conservative 1; Mismatches 98; Indels 0; Gaps 0

QY      29 TATATGCTGATTTAAGGGTGAATTTTGCTTCCTTTATACTTTTTATTTATCCCAAATT 88
Db      531 TATAAATTTATTTATATATATTTATTAATTTATTCAATTTATTTATTTTCTTAGT 590

QY      89 TTTCCTTAGCAAAATATTTCTTTGCTTAATCAATAAATTTATCAAAGAAGAAAAAACTGAAA 148
Db      591 TTATAAATAGTAATTTCTACTAAATTTAAAAAAGAAAAAAGAAAAAAGAAAA 650

QY      149 GCAACGCTTGAAAAAGGAAAGTTAGGCCCTATCGGGTATATTTTGGAAAGTTGTAAAAATAC 208
Db      651 AAAAAAATTTACATATGAAAAATGAACCTGTATATCTAAATTTATAAATATTTTAAACA 710

QY      209 TACRTGTTCTCTTTCTAA 225

```



```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/049,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match          7.3%; Score 38.8; DB 4; Length 2323;
Best Local Similarity 47.8%; Pred. No. 0.36;
Matches 109; Conservative 1; Mismatches 118; Indels 0; Gaps 0;

Qy 27 TTTATATGCTGATTTATGGGTGATTTTGGTCTCTCTTATATCTTTTATTTATCCAAA 86
Db 2291 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2232

Qy 87 TTTTCTTAAGCAAAATATTTCTTGCTAATCAATAATATCAAAAGAAAAAAACTGA 146
Db 2231 ACTCTGATAAATTTATTAATAAGAACCAAGCATCAGTTGTACATAGAAATAATCATTTG 2172

Qy 147 AAGCAACGCTTGAAAAAGAAAGTTAGCCCTATCGGTATATTTTGGAGTTGTAAAAAT 206
Db 2171 CACTAGCATGGTACTAGACAGAGGAGTTTAAATGCTATTTTTTAAGTCACAAAAT 2112

Qy 207 ACTACRTGTTCTCTTCTTAAGTCCCACTCTCTGTTTCTTTGTAGCAGG 254
Db 2111 ATATCGGTTAAGGCACCTAATACTCTGCTCATGTTTCAATAAATATG 2064

RESULT 12
US-09-212-971-3
; Sequence 3, Application US/09212971B
; Patent No. 6107041
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Teang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009002
; CURRENT APPLICATION NUMBER: US/09/212,971B
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 60/017,354
; EARLIER FILING DATE: 1996-04-26
; EARLIER APPLICATION NUMBER: 60/030,590
; EARLIER FILING DATE: 1996-11-14
; EARLIER APPLICATION NUMBER: 08/800,929
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (4623)...(4623)

```

```
; OTHER INFORMATION: n can be any nucleotide
; FEATURE:
; NAME/KEY: variation
; LOCATION: (4622)...(4622)
; OTHER INFORMATION: n can be any nucleotide
US-09-212-971-3

Query Match          7.3%; Score 38.8; DB 3; Length 5232;
Best Local Similarity 62.2%; Pred. No. 0.52;
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 34 GCTGATTTATGGGTGATTTGCTTCTTCTATATCTTTTATTTATTCCTCCAAATTTTCT 93
Db 2334 GGTTTATAGGGGCTTTTCTACTTTCTCTTTTCTTTTCTTTTCTGTTCTGTCGAATTTT 2393

Qy 94 TAAGCAAAATATTTCTTCTGCTAATCAATAAATATCAAA 131
Db 2394 TAAGTATGTATTACTTTTGTAAATCAGAATTTTAGAAA 2431

RESULT 13
US-08-800-929A-3
; Sequence 3, Application US/08800929A
; Patent No. 6133437
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF
; TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,929A
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,590
; FILING DATE: 14-NOV-1996
; APPLICATION NUMBER: 60/017,354
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 07891/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5232 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Other
```

```
; LOCATION: 1...5232
; OTHER INFORMATION: N at 4622 and 4633 can be A, G, T or C.
US-08-800-929A-3

Query Match          7.3%; Score 38.8; DB 3; Length 5232;
Best Local Similarity 62.2%; Pred. No. 0.52;
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 34 GCTGATTTATGGGTGATTTGCTTCTTCTATATCTTTTATTTATTCCTCCAAATTTTCT 93
Db 2334 GGTTTATAGGGGCTTTTCTACTTTCTCTTTTCTTTTCTTTTCTGTTCTGTCGAATTTT 2393

Qy 94 TAAGCAAAATATTTCTTCTGCTAATCAATAAATATCAAA 131
Db 2394 TAAGTATGTATTACTTTTGTAAATCAGAATTTTAGAAA 2431

RESULT 14
US-09-617-053A-3
; Sequence 3, Application US/09617053A
; Patent No. 6300492
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617,053A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; PRIOR FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (4623)...(4623)
; OTHER INFORMATION: n can be any nucleotide
; NAME/KEY: variation
; LOCATION: (4622)...(4622)
; OTHER INFORMATION: n can be any nucleotide
US-09-617-053A-3

Query Match          7.3%; Score 38.8; DB 4; Length 5232;
Best Local Similarity 62.2%; Pred. No. 0.52;
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 34 GCTGATTTATGGGTGATTTGCTTCTTCTATATCTTTTATTTATTCCTCCAAATTTTCT 93
Db 2334 GGTTTATAGGGGCTTTTCTACTTTCTCTTTTCTTTTCTTTTCTGTTCTGTCGAATTTT 2393

Qy 94 TAAGCAAAATATTTCTTCTGCTAATCAATAAATATCAAA 131
Db 2394 TAAGTATGTATTACTTTTGTAAATCAGAATTTTAGAAA 2431

RESULT 15
US-08-545-196B-10/c
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATIION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-545-196B-12

Query Match 7.2%; Score 38; DB 3; Length 1582;
Best Local Similarity 50.5%; Pred. No. 0.51;
Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 27 TTTATATGCTGATTTAATGGTGATTTTGCTTCTTATACCTTTTATTTATTTCCCAA 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1558 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1499

QY 87 TTTTCTTAAGCAATATTTCTTTGCTAATCAATAAATATCAAAAGAAAAAAACTGA 146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1498 TTTTTTTTAAAAAAAATTTAAATATTTTTTATTATATACCTTTTAAACATATAGAAGATAGA 1439

QY 147 AAGCAACGCTTGAAAAAGCAAAAGTTAGCCCTATCGGTATATTTTGGAAAGTTGTAAAAT 206
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1438 AAAAACAAGTACAAAGCAAGCATGTCACCATGATTCTGTAACTTTTGGCCAT 1379

QY 207 AC 208
|||
Db 1378 AC 1377

RESULT 17
US-09-820-005-3/c
; Sequence 3, Application US/09820005
; Patent No. 6489149

```

```

/ AFFILIATION: SPRO, NEI, ET AL
/ TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: CL001198
/ CURRENT APPLICATION NUMBER: US/09/820,005
/ CURRENT FILING DATE: 2001-03-29
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 18853
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(18853)
/ OTHER INFORMATION: n = A,T,C or G
/
/ US-09-820-005-3
/
/ Query Match 7.2%; Score 38; DB 4; Length 18853;
/ Best Local Similarity 50.5%; Pred. No. 1.5;
/ Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
/
/ QY 25 GATTATATGCTGATTATTTGGGTGATTTTGGCTTCTCTTACTTTTATTATTCCTCA 84

```

```
Db 15117 GGTGTTGTTTTGCTCTTTTTTTTTTTTTTTCAGTATGGAATCCATCTGTTCACGACC 15058
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
Qy 85 AATTTTCTTAAGCAAAATATTTCTTTGCTTAATCAATAATATCAAAAGAAAAAAACT 144
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
Db 15057 ATTGTTGAAAGACTTCTTTCTATCTCAATATATATTTGGGAGTGAAAAAAGAC 14998
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
Qy 145 GAAACAAAGCTTGAAAAAGAAAGTTAGCCCTATCGGGTATATTTTGGAAAGTTGTAAA 204
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
Db 14997 AAAAGAAACTTTGCATAATACAGGTTCCACATAAAACCAATTCTTGCTGGTGGAGG 14938
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
Qy 205 AT 206
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
Db 14937 AT 14936
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11

RESULT 18
US-09-410-464-5
; Sequence 5, Application US/09410464
; Patent No. 6395892
; GENERAL INFORMATION:
; APPLICANT: Straus et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 53375
; CURRENT APPLICATION NUMBER: US/09/410,464
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: 09/287,700
; EARLIER FILING DATE: 1999-04-06
; EARLIER APPLICATION NUMBER: 60/080,851
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 5656
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
US-09-410-464-5

Query Match 7.1%; Score 37.8; DB 4; Length 5656;
Best Local Similarity 58.4%; Pred. No. 1;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 60 TTCCTTATACCTTTTATTTATTTCCCAATTTTCTTAAGCAAAATATTTCTTCTGCTAATCAA 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4024 TTATTTAGAAATATATTAATAATATTTTATATTTTAAATTTATTTTAAATTTAA 4083
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 120 TAAATATCAAAAGAAAAAAACTGAAAGCAACGCTTGAAAAAGGAAAGTT 172
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4084 TATATTAATAATAATAAAATACTGAAAAATAATTTTTTAAAAATAATTTT 4136
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 19
US-09-149-476-66
; Sequence 66, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
```


EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576

EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 7.1%; Score 37.6; DB 4; Length 732;
Best Local Similarity 56.5%; Pred. No. 0.46;
Matches 70; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 46 GTGATTTGCTTCCTTTATACCTTTTATTTTCCCAATTTTCTTAAGCAATATT 105
DB 604 GTCAGTTTGAATCTTCTTGAAGTTTAAATGTTTATTAGGAGATTTTAAAGAAAAATAA 663
QY 106 TCTTTGCTAATCAATAAATTTATCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAAG 165
DB 664 GGTCTACAATATCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 723
QY 166 GAAA 169
DB 724 AAAA 727

RESULT 20

US-09-013-810-1/c
Sequence 1, Application US/09013810
Patent No. 6197551
GENERAL INFORMATION:
APPLICANT: Busfield, Samantha J.
TITLE OF INVENTION: No. 6197551el TANGO 80 Molecules and Uses Thereof
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/013,810
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Amy E. Mandragoras
;; REGISTRATION NUMBER: 36,207
;; REFERENCE/DOCKET NUMBER: MEI-010
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617)227-7400
;; TELEFAX: (617)742-4214
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 746 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 135..428
;;
US-09-013-810-1

Query Match 7.1%; Score 37.6; DB 3; Length 746;
Best Local Similarity 54.3%; Pred. No. 0.47; 64; Indels 0; Gaps 0;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 27 TTATATGCTGATTTATGGTGATTTGCTTCCTTCTTTATATCTTTTATTTATTTCCCAAA 86
DB 731 TTTTCTTTTAAAGCAATATTTCTTGTCTATCAATAATATCAAGAGAAAAAACTGAA 146
QY 87 TTTTCTTTAAAGCAATATTTCTTGTCTATCAATAATATCAAGAGAAAAAACTGAA 146
DB 671 TTTTCTTTAAAGCAATATTTCTTGTCTATCAATAATATCAAGAGAAAAAACTGAA 612
QY 147 AAGCAAGCTTGAAAAAGG 166
DB 611 ATGCTTTGGATGAAACATG 592

RESULT 21
US-09-221-013A-7
;; Sequence 7, Application US/09221013A
;; Patent No. 6495740
;; GENERAL INFORMATION:
;; APPLICANT: Arioli, Antonio
;; APPLICANT: Williamson, Richard E.
;; APPLICANT: Betzner, Andreas S.
;; APPLICANT: Peng, Liangcai
;; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
;; FILE REFERENCE: 96-98
;; CURRENT APPLICATION NUMBER: US/09/221,013A
;; CURRENT FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: PCT/AU97/00402
;; PRIOR FILING DATE: 1997-06-24
;; PRIOR APPLICATION NUMBER: AU P00699
;; PRIOR FILING DATE: 1996-06-27
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 7
;; LENGTH: 3828
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (239)..(3490)
US-09-221-013A-7

Query Match 7.1%; Score 37.6; DB 4; Length 3828;
Best Local Similarity 56.5%; Pred. No. 0.97;

Matches 70; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 42 ATGGTGATTTTGGTTCCTTCTTTATATCTTTTATTTATTTCCCAAAATTTTCTTAAAGCAA 101
DB 3671 ATTGTTTGGTTCCTTCTTTTACATGACTTTTAGTATTCGTTAGTTATTTATTAATACT 3730
QY 102 TATTTCTTTGCTAATAATAATATCAAGAGAAAAAACTGAAAGCAAGCTTGAA 161
DB 3731 GATTAACGATCATATATACACACTTTGTTTACAAAAAAATTTTATTAATAATAATTA 3790
QY 162 AAAG 165
DB 3791 AAAG 3794

RESULT 22
US-09-601-198-165/c
;; Sequence 165, Application US/09601198
;; Patent No. 6531583
;; GENERAL INFORMATION:
;; APPLICANT: Cassell, Gail H.
;; APPLICANT: Chen, Jennifer S.
;; APPLICANT: Glass, John I.
;; APPLICANT: Heiner, Cheryl R.
;; APPLICANT: Lefkowitz, Elliot
;; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREA PLASMA
;; FILE REFERENCE: UAB-13452/22
;; CURRENT APPLICATION NUMBER: US/09/601,198
;; CURRENT FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/073,189
;; PRIOR FILING DATE: 1998-01-30
;; NUMBER OF SEQ ID NOS: 181
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 165
;; LENGTH: 4344
;; TYPE: DNA
;; ORGANISM: Ureaplasma urealyticum
US-09-601-198-165

Query Match 7.1%; Score 37.6; DB 4; Length 4344;
Best Local Similarity 54.3%; Pred. No. 1;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 28 TTATATGCTGATTTATGGTGATTTGCTTCCTTCTTTATATCTTTTATTTATTTCCCAAT 87
DB 3278 TTAATTGATGGTGAATTTGTTGTAATTTCTTTTAAATATTCATCTTTATGATCAACT 3219
QY 88 TTTTCTTAAAGCAATATTTCTTGTCTAATCAATAATATCAAGAGAAAAAACTGAA 147
DB 3218 AAATCAATTAACACTTTCATCAATAATCTTGATTAATTTTTCATCACTTATAATAATTA 3159
QY 148 AGCAAGCTTGAAAAAGG 167
DB 3158 TATGATCTTTTAAAGATGA 3139

RESULT 23
US-08-916-421B-1
;; Sequence 1, Application US/08916421B
;; Patent No. 6503729
;; GENERAL INFORMATION:
;; APPLICANT: Bult et al.
;; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
;; FILE REFERENCE: PB275
;; CURRENT APPLICATION NUMBER: US/08/916,421B
;; CURRENT FILING DATE: 1997-08-22
;; PRIOR APPLICATION NUMBER: US 60/024,428
;; PRIOR FILING DATE: 1996-08-22
;; NUMBER OF SEQ ID NOS: 3

```

/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1664976
/ TYPE: DNA
/ ORGANISM: Methanococcus jannaschii
/ FEATURE:
/   NAME/KEY: misc feature
/   LOCATION: (28222)..(28222)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (28257)..(28258)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (84773)..(84773)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (84808)..(84808)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (84812)..(84812)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (98120)..(98120)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (98159)..(98159)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (98239)..(98239)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (98266)..(98266)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (98343)..(98343)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (103998)..(103998)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (148948)..(148948)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (163385)..(163385)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (191989)..(191989)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (191995)..(191995)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (231980)..(231980)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (234187)..(234187)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (234220)..(234220)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (234814)..(234814)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (309398)..(309398)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (309418)..(309418)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature
/   LOCATION: (312837)..(312837)
/   OTHER INFORMATION: n equals a, t, c, or g
/   NAME/KEY: misc feature

```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match
Best Local Similarity 7.1%; Score 37.6; DB 4; Length 1664976;
Matches 91; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 24 TGATTATATCTGATTTATGGTGATTTTGGTTCCTTTATACATTTTATTTATTTCC 83
Db 1332304 TAAATCAACATCTACTCATTTAGTTTATATAATTCCTTGCNAACATTTTAAAACTCC 1332363
Qy 84 AAATTTTCTTAAGCAATATTTCTTGTCTATCAATAAATATCAAAAGAAAAAAAC 143
Db 1332364 ACTGTGTTTTAGGATTAATTTCTTTTGTAAATTAAGCCTTAACATTAACAT. 1332423
Qy 144 TGAAGCAACGCTTGAAAAAGGAAAGTTAGCCCTATCGGTATATTTTGGAAAGTTGPA 203
Db 1332424 TGAGTAGTATATCTTGAACAGCAAAATCATAAAATTCGCTATTATAAGATTTCTGA 1332483

RESULT 24
US-09-369-247-23/c
Sequence 23, Application US/09369247
Patent No. 6569992
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 44 Human Secreted Proteins
FILE REFERENCE: P2024P1
CURRENT APPLICATION NUMBER: US/09/369,247
EARLIER FILING DATE: 1999-08-05
EARLIER APPLICATION NUMBER: 60/074,118
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,157
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,137
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,341
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,141
EARLIER FILING DATE: 1998-02-09
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 1492
TYPE: DNA
ORGANISM: Homo sapiens
US-09-369-247-23

Query Match
Best Local Similarity 7.0%; Score 37.2; DB 4; Length 1492;
Matches 57; Conservative 2; Mismatches 35; Indels 0; Gaps 0;
Qy 39 TTTATGGGTGATTTTGGTTCCTTTATACATTTTATTTATTTCCCAATTTTCTTAAGC 98
Db 1484 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1425
Qy 99 AAATATTTCTTGTCTAATCAATAAATATCAAAA 132

Db 1424 TTTTATTTTGTGAAAAACCAATAATTTATCAAAA 1391
RESULT 25
US-07-867-106-2
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Peeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match
Best Local Similarity 7.0%; Score 37.2; DB 1; Length 5852;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
Qy 27 TTTATATGCTGATTTATGGGTGATTTTGGTTCCTTCTTTATACATTTTATTTATTTCCCAAA 86
Db 5594 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5653
Qy 87 TTTTCTTACCAATATTTCTTGTCTAATCAATAAATATCAAAAGAAAAAACTGA 146
Db 5654 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5713
Qy 147 AA 148
Db 5714 AA 5715

RESULT 26

US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-545-196B-10

Query Match 7.0%; Score 37; DB 3; Length 1582;
Best Local Similarity 56.0%; Pred. No. 0.95;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 45 GGTGATTTTGGCTTCCTCTTATCTTTTATTTATTTCCCAAAATTTTCTTAAGCAAAATAT 104
Db 1420 GTTCATTGTACTGTTTTTCTCTCTATATGTTTAAAGTATATATAAAATATTT 1479
QY 105 TTCTTTGCTAAATCAATAATTTCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 164
Db 1480 AATTTTTTTTAAAAAATAATTTCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 1539
QY 165 GGAAA 169
Db 1540 AAAAA 1544

RESULT 27

US-08-545-196B-12
; Sequence 12, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH

STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,196B
FILING DATE: 19-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FARACI, C. J.
REGISTRATION NUMBER: 32,350
REFERENCE/DOCKET NUMBER: 2121-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-545-196B-12

Query Match 7.0%; Score 37; DB 3; Length 1582;
Best Local Similarity 56.0%; Pred. No. 0.95;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 45 GGTGATTTTGGCTTCCTCTTATCTTTTATTTATTTCCCAAAATTTTCTTAAGCAAAATAT 104
Db 1420 GTTCATTGTACTGTTTTTCTCTCTATATGTTTAAAGTATATATAAAATATTT 1479
QY 105 TTCTTTGCTAAATCAATAATTTCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 164
Db 1480 AATTTTTTTTAAAAAATAATTTCAAAAGAAAAAACTGAAAGCAACGCTTGAAAAAA 1539
QY 165 GGAAA 169
Db 1540 AAAAA 1544

RESULT 28

US-08-747-221B-36
; Sequence 36, Application US/08747221B
; Patent No. 6063610
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heeka Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,221B
; FILING DATE: No. 6063610ember 12, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: FC-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 2007 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 11..1594
US-08-747-221B-36

Query Match 7.0%; Score 37; DB 3; Length 2007;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 24 TGATTATATGCTGATTTGGTGATTTGCTTCCTTCTTTATATCTTTTATTTATTTCC 83
Db 1893 TTATTTATATTTGATATATTACCATCTTTGATCATATTTGCTTTTATTTTTCATTTT 1942
QY 84 AAATTTTCTTAAGCAATATTTCTTGTCTATCAATAATATCAAAAGAAAAAAC 143
Db 1943 TTTTATTTCAATATATTTGTTTTTATATATATATATATATATATATATATATAT 2002
QY 144 TGA 148
Db 2003 AAAAA 2007

RESULT 29
US-08-747-221B-38/c
Sequence 38, Application US/08747221B
Patent No. 6063610
GENERAL INFORMATION:
APPLICANT: Silver, Gary W.
APPLICANT: Wisniewski, Nancy
TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,221B
FILING DATE: No. 6063610ember 12, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: FC-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 2007 nucleotides
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-747-221B-38
Query Match 7.0%; Score 37; DB 3; Length 2007;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 24 TGATTATATGCTGATTTGGTGATTTGCTTCCTTCTTTATATCTTTTATTTATTTCC 83
Db 125 TTATTTATATTTGATATATTACCATCTTTGATCATATTTGCTTTTATTTTTCATTTT 66
QY 84 AAATTTTCTTAAGCAATATTTCTTGTCTATCAATAATATCAAAAGAAAAAAC 143
Db 65 TTTTATTTCAATATATTTGTTTTTATATATATATATATATATATATATATATATAT 6
QY 144 TGA 148
Db 5 AAAAA 1
RESULT 30
US-09-005-051-36
Sequence 36, Application US/09005051
Patent No. 6291222
GENERAL INFORMATION:
APPLICANT: Silver, Gary W.
APPLICANT: Wisniewski, Nancy
TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,051
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/747,221
FILING DATE: No. 6291222ember 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: FC-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 2007 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 11..1594
US-09-005-051-36
Query Match 7.0%; Score 37; DB 3; Length 2007;
Best Local Similarity 56.0%; Pred. No. 1.1;

APPLICANT: Artiushin, Sergey
APPLICANT: Stipkovits, Laslo
APPLICANT: Minion, F. Chris

[illegible]

148 ACACACCC11GAAAAAGGAAA 188
 | | | | | | | | | | | | | |
 Db 578 AAAACCGAAGGAATTAGTAA 598

 RESULT 33
 US-09-234-245-1
 ; Sequence 1, Application US/09234245
 ; Patent No. 6509457
 ; GENERAL INFORMATION:
 ; APPLICANT: Thomas Jefferson University
 ; APPLICANT: Raphael Rubin
 ; APPLICANT: Manorama Tewari


```
; CURRENT APPLICATION NUMBER: US/09/453,323
; CURRENT FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: 60/110,676
; PRIOR FILING DATE: 1998-12-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-09-453-323-1

Query Match
Best Local Similarity 6.8%; Score 36; DB 4; Length 1507;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 46 GTGATTTTGCCTTCCTTATACATTTTATTTCCCAATTTTCTTAAGCAATATT 105
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1357 GTGATTTGATTTGAGCTGTATTCGCTGTATTCCTCAATTCCTCCCTAAGCAAGATAT 1416
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 106 TCTTTGCTAATCAATAAATTATCAAAAGAAAAAACTGAAAGCAAGCTTGAAAAAG 165
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1417 TAGCAGATGATAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1476
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 166 GAAA 169
Db |||||
Qy 1477 AAAA 1480
Db |||||

RESULT 40
US-08-821-994-59
; Sequence 59, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier RP
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821,994A
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 1577
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-59

Query Match
Best Local Similarity 6.8%; Score 36; DB 3; Length 1577;
Matches 78; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 22 GCTGATTTATATGCTGATTTATGGTGATTTTGGCTTCCTTCTTATATCTTTTATTC 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1426 GATGATTTACTTTATAGCTGTCTTGTGATATGATATATTAGTCTCTTATTTGGATGATA 1485
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 82 CCAAAATTTTCTTAAGCAAAATTTTCTTTGCTTAATCAATAAATTAATCAAGAAAAA 141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1486 CAACTTTTGAATCAATAAAGTTTACTTGCAGGACACATAAAGAAAAAAGAAAAA 1545
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 142 ACTGAAGCAAGCTTGAAAAAGAAA 169
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1546 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1573
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 41
US-09-293-322C-8
; Sequence 8, Application US/09293322C
```

```
; Patent No. 6232110
; GENERAL INFORMATION:
; APPLICANT: Pallas, David C
; APPLICANT: Du, Xianxing
; TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,
; Patent No. 6232110
; TITLE OF INVENTION: Recombinant DNA Molecules and Methods
; FILE REFERENCE: 105-97
; CURRENT APPLICATION NUMBER: US/09/293,322C
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 60/082,202
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2409)
; OTHER INFORMATION: N is A, T, G or C.
US-09-293-322C-8

Query Match
Best Local Similarity 6.8%; Score 36; DB 3; Length 2409;
Matches 60; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 70 TTTTATTATTTCCCAATTTTCTTAAGCAATATTTCTTGGCTAATCAATAAATATCA 129
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2300 TCTTTGTTATTTATGATCTTCTTTAAGAAAAAATAATATCTCCCAACCTTTTAAAAA 2359
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 130 AAAGAAAAAAGCTGAAAGCAAGCTTGAAAAAAGGAAA 169
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2360 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2399
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 42
US-09-839-497A-8
; Sequence 8, Application US/09839497A
; Patent No. 6528295
; GENERAL INFORMATION:
; APPLICANT: Pallas, David C.
; APPLICANT: Du, Xianxing
; TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,
; Patent No. 6528295
; TITLE OF INVENTION: Recombinant DNA Molecules and Methods
; FILE REFERENCE: Docket No. 6528295 105-97A
; CURRENT APPLICATION NUMBER: US/09/839,497A
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/082,202
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/293,322
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2409)
; OTHER INFORMATION: N is A, T, G or C.
US-09-839-497A-8

Query Match
Best Local Similarity 6.8%; Score 36; DB 4; Length 2409;
Matches 60; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 70 TTTTATTATTTCCCAATTTTCTTAAGCAATATTTCTTGGCTAATCAATAAATATCA 129
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2300 TCTTTGTTATTTATGATCTTCTTTAAGAAAAAATAATATCTCCCAACCTTTTAAAAA 2359
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE

; LOCATION: (35)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-227-357-76

Query Match 6.7%; Score 35.6; DB 4; Length 519;
Best Local Similarity 59.0%; Pred. No. 1.4;
Matches 59; Conservative 1; Mismatches 40; Indels 0; Gaps 0;

Qy 46 GTGATTTTGGCTTCCTTTTATATCTTTTATATCCCAAAATTTTCTTAAGCAAATATT 105
Db 497 GAGTTTTTTTTTWTYTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAACATTATT 438

Qy 106 TCCTTGCTAATCAATAAATTATCAAAAGAAAAAAACTG 145
Db 437 TATCTACTGTACAAATAATTTTACATCATCATGCTGCAACTG 398

Search completed: February 14, 2004, 15:19:13
Job time : 81 secs

THIS PAGE BLANK (USPTO)